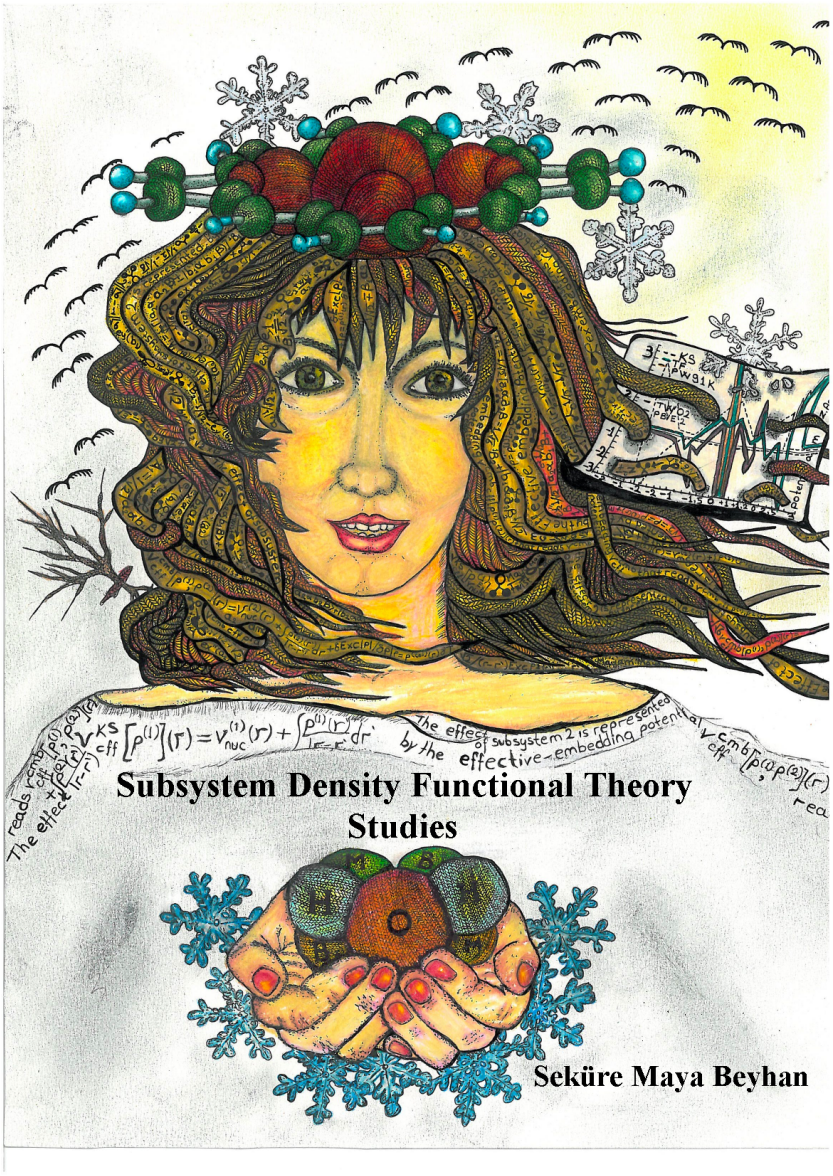
Subsystem Density Functional Theory Studies

Beyhan, S.M.

2012

document version
Publisher's PDF, also known as Version of record



Subsystem Density Functional Theory Studies

Seküre Maya Beyhan

VRIJE UNIVERSITEIT

Subsystem Density Functional Theory Studies

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. L.M. Bouter,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Exacte Wetenschappen
op woensdag 16 januari 2013 om 15.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Maya Beyhan

geboren te Ankara, Turkije

promotor: prof.dr. L. Visscher

VRIJE UNIVERSITEIT

Subsystem Density Functional Theory Studies

Maya Beyhan

Cover painting by Nevin Beyhan

Contents

Chapter 1

General Introduction

1.1 Introduction

Quantum chemistry is a theory which mathematically describes the fundamental behaviour of atoms and molecules to address various problems in chemistry. By employing quantum chemical methods, one can calculate molecular properties, which directly depend on the electron density distribution such as electric moments, polarizabilities, shielding constants in nuclear magnetic resonance (NMR) or excited states and structures with unusual electronic effects. Such fundamental calculations can clarify many important questions that cannot be addressed by experiments and empirical approaches only.

Quantum chemical methods are based on the postulates of quantum mechanics. These methods, which solve the Schrödinger equation, can be divided into the wave function based methods and density-functional theory (DFT). The wave function based methods[1] have the advantage of systematically approaching the exact solution of Schrödinger equation but the required computer time increases substantially with the accuracy of the wave function method employed.

On the other hand, DFT[2] avoids the calculation of the many-electron wave function by replacing it with the simpler electron density, that can be represented in terms of the density of an auxiliary non-interacting system in the Kohn-Sham (KS) approach. This approach makes the computational cost is in most cases lower than that of the wave function based methods. However, the exact functional form of the exchange-correlation energy, which is the key ingredient of KS-DFT is not known and needs to be approximated. An improvement, thus, can

be reached with each successive better approximation to the exchange-correlation energy. Provided that a suitable exchange-correlation functional is employed, the KS method represents one of the most successful quantum chemical methods. A review of DFT with a focus on the KS method can be found in Chapter 2.

Even though the computational cost of DFT is so much cheaper than that of wave function based methods, it is still prohibitively expensive to study phenomena on large length and time scales as they are typically required for large systems such as biologically relevant ones.

However, in most cases, it is neither desirable nor necessary to describe a system in its entirety because most properties of interests are local in nature. Thus, the total system can be divided into a localized electronic event in an active site (e.g. interaction between an enzyme and its substrate) and an interacting larger environment which influences this active site (e.g. protein and solvent). A reasonable approach for modelling biological systems is, thus, the use of subsystem methods that focus the attention on a particular region of interest, like an active site, treating the interacting environment at a different level.[3–6]

A very popular subsystem method is the quantum mechanics/molecular mechanics (QM/MM) method, as pioneered by Warshel and co-workers,[7,8] on which there has been a considerable amount of recent reviews documenting its developments and applications.[9–20] In the QM/MM scheme, a quantum chemical method, either DFT or wave function based, is used to describe the chemically active region which is combined with efficient force-field based molecular mechanics (MM) methods to treat the interacting surroundings. This interaction between the subsystem of interest and the environment is modeled by an interaction potential that depends on the careful calibration of a large number of parameters against experimental reference data.

Another type of subsystem approach is the combined QM/QM schemes[21–23] where different quantum chemical methods are used for different subsystems. For instance, the chemically active region is treated with a highly accurate wave function based method, whereas a DFT description of the interacting surrounding is employed. Employing the same method for both regions is also possible such as in a DFT-in-DFT embedding scheme. However, when wave function based methods are employed in the QM/QM method, it is difficult to describe the interaction between the subsystem of interest and the environment. This is because such a description calls for a consistent theory on how to partition the wave function into subsystems and how to match them at the boundaries, which brings along issues such as the use of additional approximations.[24,25]

On the other hand, such a partitioning of the wave function is not needed within DFT, since it avoids the calculation of the many-electron wave function by replacing it with the electron density. Furthermore, unlike the QM/MM method, no parameterization is needed to model the interaction potential for the subsystem DFT formalism as it was initially proposed by Cortona,[26] which offers a plausible alternative to the conventional KS method. Subsystem DFT founds the so-called frozen density embedding (FDE) method within DFT as proposed by Wesolowski and Warshel.[21] In FDE, the electron density can be partitioned and that of the subsystem of interest is determined in the presence of an effective embedding potential[21] due to the electron density of the environment. Subsystem DFT will be discussed in greater detail in Chapter 3.

Subsystem DFT is an efficient alternative to the conventional KS method as it allows an accurate description of the interaction between the subsystem of interest and the environment with shorter computer time. Moreover, unlike the QM/MM method, no parameterization is needed to model this interaction. Owing to these advantages, subsystem DFT is a scheme that shows indication of future excellence to study large systems where the balance of efficiency and accuracy is of great importance. Jacob[27] has extended the applicability of subsystem DFT and explored some of its possibilities. This thesis attempts to identify in detail the current challenges in the subsystem DFT scheme and sets out to further explore its potential.

1.2 Overview of This Thesis

This thesis investigates the present shortcomings, and their reasons, of subsystem DFT by a thorough analysis of the effective embedding potential. It gives a particular focus to the behavior of the effective embedding potential in the presence of which the electron density of the subsystem of interest is determined. Moreover, it further evaluates the applicability of subsystem DFT.

Chapter 2 gives a review of DFT with a special focus on the KS method. The subsystem DFT scheme, which is an efficient alternative to the KS method, is introduced in Chapter 3.

In Chapter 4, a study where the performance of the effective embedding potentials produced by the currently available approximations used within the FDE scheme have been assessed for describing the weak covalent bond found in the noble gas atom bound to a gold fluoride molecule, is presented.

Chapter 5 presents the construction of an accurate effective embedding potential for a given electron density of the subsystem of interest, employed on diatomic molecules with covalent bonding.

Chapter 6 discusses the convergence behaviour of total interaction energy obtained with the FDE scheme, where we adopted a similar approach to the KS bond formation analysis,[28] by rewriting the total interaction energy with respect to promotion energy of individual subsystems and interaction energy between these. We have split up these energy terms into their corresponding contributions and analysed each for different freeze-and-thaw cycles until convergence, for a representative set of hydrogen-bonded and stacked base pairs as well as for interacting amino acid pairs found in the protein rubredoxin.

Finally, Chapter 7 is adapted from a work,[SMB1] where no subsystem DFT scheme was used, however, an explanation for the anomalous fluorescence behavior of isoflavones is presented, where the solvent effects have been analyzed.

Chapter 2

Review of Density Functional Theory

In this Chapter a review of the density-functional theory (DFT) is given. Chapter 2.1 reviews the elementary quantum mechanics where the principles of DFT stem out. Chapter 2.2 expounds DFT. The Hohenberg-Kohn theorem and the Kohn-Sham method are introduced and the components of the total energy functional, particularly the kinetic-energy and the exchange-correlation energy are discussed. This is followed by an overview of the orbital-free DFT scheme. A particular focus is given on the differences between the Kohn-Sham method and orbital-free DFT in handling the kinetic-energy, which is a fundamental issue in the frozen-density embedding method.

2.1 Quantum Mechanics

The behavior of very small particles such as the electrons and nuclei of atoms and molecules is described by the theory of quantum mechanics.[29]

In quantum mechanics, stationary states are states of constant energy that experience no time-dependent external forces. In most cases, one is concerned with such stationary states, e.g. the stationary state with the lowest energy, the ground state of the system. For stationary systems, therefore, the wave function $\Psi(\boldsymbol{r}_1, s_1, \ldots, \boldsymbol{r}_N, s_N)$ that contains all possible information is a function of particles' coordinates only, where $\boldsymbol{r}_i$ are the continuous spatial coordinates and s_i the discrete spin coordinates.

The probability density is given by $|\Psi|^2$ such that $|\Psi(\boldsymbol{r}_N, s_N)|^2 d\boldsymbol{r}_N$ is the probability of finding a particle with position coordinates between $\boldsymbol{r}_N$ and $\boldsymbol{r}_N + d\boldsymbol{r}_N$ and spin coordinates equal to s_N. The wave function Ψ is normalized according to

$$\int_{-\infty}^{\infty} |\Psi|^2 d\boldsymbol{r}_1 ds_1, \ldots, d\boldsymbol{r}_N ds_N = 1. \tag{2.1}$$

$|\Psi|^2$ does not change with time for systems at stationary states. They can, therefore, be obtained, for an isolated N-electron atomic or molecular system in the nonrelativistic approximation, by the time-independent Schrödinger's equation which reads

$$\hat{H}\Psi = E\Psi \tag{2.2}$$

where E is the electronic energy and $\hat{H}$ is the Hamiltonian operator,

$$\hat{H} = \sum_{i=1}^{N}(-\frac{1}{2}\nabla_i^2) + \sum_{i=1}^{N} v(\boldsymbol{r}_i) + \sum_{i<j}^{N} \frac{1}{\boldsymbol{r}_{ij}} \tag{2.3}$$

in which

$$v(\boldsymbol{r}_i) = -\sum_{\alpha} \frac{Z_\alpha}{\boldsymbol{r}_{i\alpha}} \tag{2.4}$$

is the external potential acting on electron i, the potential due to nuclei of charges Z_α. Atomic units are employed here and throughout this thesis. Eq. (2.3) can be written more compactly as

$$\hat{H} = \hat{T} + \hat{V}_{ne} + \hat{V}_{ee} \tag{2.5}$$

where

$$\hat{T} = \sum_{i=1}^{N}(-\frac{1}{2}\nabla_i^2) \tag{2.6}$$

is the kinetic energy operator,

$$\hat{V}_{ne} = \sum_{i=1}^{N} v(\boldsymbol{r}_i) \tag{2.7}$$

is the electron-nucleus attraction energy operator, and

$$\hat{V}_{ee} = \sum_{i<j}^{N} \frac{1}{\boldsymbol{r}_{ij}} \tag{2.8}$$

is the electron-electron repulsion energy operator.

Eq. (2.2) contains two unknowns, the allowed energies E and wave functions Ψ. To solve for two unknowns, appropriate boundary conditions should be imposed on Ψ. It has to be normalizable and well-behaved everywhere, i.e. smooth, continuous and differentiable. In addition, because electrons are fermions, Ψ must be antisymmetric with respect to interchange of space and spin coordinates of any two electrons. Only certain values of E allow Ψ to meet these boundary conditions.

The eigen functions Ψ_i with corresponding energy eigen values E_i constitute an infinite set of independent solutions of Eq. (2.2) for a given system. Ψ_i forms a complete and orhonormal set, as

$$\int \Psi^* \Psi_i d\boldsymbol{r}_N ds_N = \langle \Psi | \Psi_i \rangle = 1, \tag{2.9}$$

where the star denotes the complex conjugate. The wave function Ψ is complex, but the real probability density $|\Psi|^2$ is the quantity that is experimentally observable, and is given by the product of Ψ with its complex conjugate.

Expectation values of other observables are given by

$$\langle \hat{O} \rangle = \frac{\int \Psi^* \hat{O} \Psi d\boldsymbol{r} ds}{\int \Psi^* \Psi d\boldsymbol{r} ds} = \frac{\langle \Psi | \hat{O} | \Psi \rangle}{\langle \Psi | \Psi \rangle}, \tag{2.10}$$

where $\hat{O}$ is the Hermitian linear operator for the observable O. Eigenvalues of $\hat{O}$ are the possible results of a measurement of O where individual measurements give particular eigenvalues of $\hat{O}$. For example, the expectation value of the kinetic energy is given by

$$T[\Psi] = \langle \hat{T} \rangle = \int \Psi^* \hat{T} \Psi d\boldsymbol{r} ds, \tag{2.11}$$

where the square bracket signifies that T is a functional of Ψ.

2.1.1 Variation Method

Due to the electron-electron interaction term in the Hamiltonian in Eq. (2.2), the Schrödinger's equation for many-electron atoms and molecules cannot be solved exactly. Therefore, approximate methods of solution must be employed. Among such methods is the Variation Method, which allows to approximate the ground-state energy of a system without solving the Schrödinger's equation exactly.

For a given system with the wave function Ψ, the average of many measurements (see Eq. (2.10) and the corresponding discussion for the interpretation of measurements) of the energy is given by

$$E[\Psi] = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle}, \tag{2.12}$$

where

$$\langle \Psi | \hat{H} | \Psi \rangle = \int \Psi^* \hat{H} \Psi d\boldsymbol{r} ds. \tag{2.13}$$

Each particular measurement of the energy gives one of the eigenvalues of $\hat{H}$, which leads to

$$E[\Psi] \geq E_0, \tag{2.14}$$

where E_0 is the ground-state energy. The ground-state wave function Ψ_0 and energy $E[\Psi_0] = E_0$ can be obtained by minimization of the functional $E[\Psi]$ with respect to all allowed N-electron wave functions

$$E_0 = \min_{\Psi} E[\Psi]. \tag{2.15}$$

A good choice for the trial wave function Ψ determines the success of the variation method.

2.1.2 Hartree-Fock Approximation

In the Hartree-Fock method,[30,31] the wave function Ψ is approximated as an antisymmetrized product of N orthonormal one-electron functions, i.e. orbitals. These are spin orbitals $\psi_i(\boldsymbol{r}, s)$, each a product of a spatial orbital $\phi_k(\boldsymbol{r})$ and a spin function $\sigma(s) = \alpha(s)$ or $\beta(s)$. This provides the most simple wave function that satisfies the Pauli principle: the Slater determinant

$$\Psi_{HF} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\boldsymbol{r}_1, s_1) & \psi_2(\boldsymbol{r}_1, s_1) & \cdots & \psi_N(\boldsymbol{r}_1, s_1) \\ \psi_1(\boldsymbol{r}_2, s_2) & \psi_2(\boldsymbol{r}_2, s_2) & \cdots & \psi_N(\boldsymbol{r}_2, s_2) \\ \vdots & \vdots & & \vdots \\ \psi_1(\boldsymbol{r}_N, s_N) & \psi_2(\boldsymbol{r}_N, s_N) & \cdots & \psi_N(\boldsymbol{r}_N, s_N) \end{vmatrix}. \tag{2.16}$$

Eq. (2.12) is, thus, minimized for this determinantal form of Ψ by finding the orthonormal orbitals ψ_i that minimize the expectation value of the Hamiltonian.

However, the single determinantal description does never give the exact wave function for a system of many interacting electrons. Therefore, the ground-state

energy provided by the Hartree-Fock method is only an approximate ground-state energy, which is an upper bound to the correct ground-state energy, and the wave function found is an approximate one. For a higher accuracy, one should employ a linear combination of many Slater determinants, and thus a larger search space, to minimize Eq. (2.12). This leads to wave function based electron correlation methods.

2.1.3 Electron Density

For a given state, the number of electrons in a unit volume is the electron density of that state, designated as $\rho(\boldsymbol{r})$. The electron density is related to the wave function Ψ as,

$$\rho(\boldsymbol{r}) = N \int \cdots \int |\Psi(\boldsymbol{r}, s, \boldsymbol{r}_2, s_2, \cdots, \boldsymbol{r}_N, s_N)|^2 ds d\boldsymbol{r}_2 ds_2 \cdots d\boldsymbol{r}_N ds_N. \quad (2.17)$$

The integration runs over the coordinates of electrons 2 through N, because to obtain the probability of finding electron 1 in a unit volume, one has to consider all possible locations for electrons $2, 3, \cdots, N$. Furthermore, due to the indistinguishability of electrons, probabilities are equal for different electrons. Thus, for a system of N electrons, the probability of finding an electron in a volume element is N times that of electron 1.

$\rho(\boldsymbol{r})$ is a simple function of three coordinates, $x, y,$ and z, that integrates to the total number of electrons,

$$\int \rho(\boldsymbol{r}) d\boldsymbol{r} = N. \quad (2.18)$$

The electron density decreases monotonically away from the nucleus for an atom in its ground-state. For molecules, on the other hand, buildups of density in the bonding region are seen. At the nucleus, the electron density has a finite value.

Electron density $\rho(\boldsymbol{r})$ and its associated calculation scheme will be the topic of following chapters.

2.2 Density Functional Theory

2.2.1 Hohenberg-Kohn Theorem

The ground-state wave function can be determined by the use of variational principle combined with the Hartee-Fock method or the linear mixing of many Slater determinants. However, because the N-electron wave function $\Psi(\boldsymbol{r}_1, s_1, \ldots, \boldsymbol{r}_N, s_N)$ depends on $3N$ spatial and N spin coordinates, one often has to include millions of Slater determinants in order to obtain an accurate description of the wave function. Even with a more compact representation of the wave function provided by coupled cluster[32] or density matrix renormalization group (for a review see Ref. 33), this is still a very complicated and demanding method.

The density-functional theory (DFT) allows one to simplify this problem by using the electron density $\rho(\boldsymbol{r})$, which is a function of only three coordinates, to calculate the energy and other properties of atoms and molecules.

The first Hohenberg-Kohn theorem[34] proves that the electron density $\rho(\boldsymbol{r})$ contains all information that is needed to calculate the electron density. It states that the external potential v_{ext} is determined by the electron density $\rho(\boldsymbol{r})$, which determines the number of electrons. The reason is that once the external potential and the number of electrons are known, the full Hamiltonian is defined. This means that $\rho(\boldsymbol{r})$ also determines the ground-state wave function Ψ_0 and all other electronic properties of the system, for example the kinetic energy $T[\rho]$, the potential energy $V[\rho]$, and the total energy $E[\rho]$. For a given external potential v_{ext} we may, therefore, write the following density functional

$$E_v[\rho] = \int \rho(\boldsymbol{r}) v_{\text{ext}}(\boldsymbol{r}) d\boldsymbol{r} + T[\rho] + V_{ee}[\rho]. \tag{2.19}$$

This one-to-one mapping between the electron density $\rho(\boldsymbol{r})$ and the ground-state wave function Ψ_0 is essential. It leads to a significant simplifications as it provides equations that are easier to solve.

The second Hohenberg-Kohn theorem[34] provides the energy variational principle. For a trial density $\rho(\boldsymbol{r})$ such that $\rho(\boldsymbol{r}) \geq 0$ everywhere and $\int \rho(\boldsymbol{r}) d\boldsymbol{r} = N$, the relation between the ground-state energy E_0 and the energy functional $E_v[\rho]$ in Eq. (2.19) is as the following

$$E_v[\rho] \geq E_0. \tag{2.20}$$

One may note the analogy to the variational principle for wave functions in Eq.

(2.14). The ground-state energy E_0 and density ρ_0 can be calculated by minimising the total energy functional, similar to Eq. (2.15),

$$E_0 = \min_{\rho} E[\rho]. \tag{2.21}$$

However, there is the remaining problem concerning the calculation of the interacting kinetic-energy $T[\rho]$ and the non-classical part of the electron-electron repulsion energy $V_{ee}[\rho]$. We may write

$$V_{ee}[\rho] = J[\rho] + V_{ee}^{\text{nonclassical}}[\rho], \tag{2.22}$$

where $J[\rho]$ is the Coulomb potential energy given by

$$J[\rho] = \frac{1}{2}\int\int \frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r'})}{|\boldsymbol{r'}-\boldsymbol{r}|} d\boldsymbol{r} d\boldsymbol{r'}, \tag{2.23}$$

which can be calculated explicitly in terms of the electron density. On the other hand, the explicit form of the density functionals of the interacting kinetic-energy $T[\rho]$ and the non-classical electron-electron repulsion energy $V_{ee}^{\text{nonclassical}}$ are not known. The interacting kinetic-energy functional reads

$$T[\rho] = -\frac{1}{2}\left\langle \Psi \middle| \sum_{i=1}^{N} \nabla_i^2 \middle| \Psi \right\rangle, \tag{2.24}$$

and the non-classical electron-electron repulsion is given by

$$V_{ee}^{\text{nonclassical}}[\rho] = \sum_{i=1}^{N}\sum_{j=i+1}^{N} \left\langle \Psi \middle| \frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \middle| \Psi \right\rangle - J[\rho]. \tag{2.25}$$

2.2.2 Kohn-Sham Method

Kohn and Sham[35] invented an indirect approach to simplify the calculation of the kinetic-energy functional $T[\rho]$, the Kohn-Sham (KS) method. According to this method, a non-interacting reference system are introduced into the problem so that the kinetic-energy can be appoximated to high accuracy and the remaining correction is handled separately. In order to give a unique value to this approximate kinetic-energy by having a unique decomposition in terms of orbitals, for a given ρ, Kohn and Sham established a noninteracting reference system with the external potential v_s and the Hamiltonian

$$\hat{H}_s = \sum_{i}^{N}\Big(-\frac{1}{2}\nabla_i^2\Big) + \sum_{i}^{N} v_s(\boldsymbol{r}), \tag{2.26}$$

which does not contain electron-electron repulsion terms, and for which the ground-state electron density is exactly ρ. This noninteracting reference system has an exact determinantal ground-state wave function

$$\Psi_s = \frac{1}{\sqrt{N!}}\det[\psi_1\psi_2\cdots\psi_{\mathrm{N}}], \tag{2.27}$$

where ψ_i are the N lowest eigenstates of the one-electron Hamiltonian $\hat{h}_s$:

$$\hat{h}_s\psi_i = \Big[-\frac{1}{2}\nabla^2 + v_s(\boldsymbol{r})\Big]\psi_i = \varepsilon_i\psi_i. \tag{2.28}$$

The noninteracting kinetic-energy $T_s[\rho]$ is given by

$$T_s[\rho] = \sum_{i=1}^{N}\left\langle\psi_i\middle|-\frac{1}{2}\nabla_i^2\middle|\psi_i\right\rangle, \tag{2.29}$$

which is defined for any noninteracting v-representable density, i.e. for any density that is a ground-state density of a well-defined noninteracting reference system.

By introducing this noninteracting kinetic-energy $T_s[\rho]$, the total energy functional in Eq. (2.19) is rewritten as

$$E_v[\rho] = \int\rho(\boldsymbol{r})v_{\mathrm{nuc}}(\boldsymbol{r})d\boldsymbol{r} + T_s[\rho] + J[\rho] + E_{\mathrm{xc}}[\rho], \tag{2.30}$$

where $\int\rho(\boldsymbol{r})v_{\mathrm{nuc}}(\boldsymbol{r})d\boldsymbol{r}$ is the electrostatic attraction of the electrons and the nuclei, and

$$E_{\mathrm{xc}}[\rho] = T[\rho] - T_s[\rho] + V_{ee}[\rho] - J[\rho]. \tag{2.31}$$

The quantity $E_{\mathrm{xc}}[\rho]$ is called the exchange-correlation energy and it contains the difference between T and T_s plus $V_{ee}^{\mathrm{nonclassical}}$ as defined in Eq. (2.22). The Kohn-Sham treatment follows by the introduction of the quantity $v_{\mathrm{eff}}^{\mathrm{KS}}$, the KS effective potential, defined by

$$v_{\mathrm{eff}}^{\mathrm{KS}} = v_{\mathrm{nuc}}(\boldsymbol{r}) + \int\frac{\rho(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r'}|}d\boldsymbol{r'} + v_{\mathrm{xc}}(\boldsymbol{r}), \tag{2.32}$$

with the exchange-correlation potential

$$v_{\mathrm{xc}}(\boldsymbol{r}) = \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta\rho(\boldsymbol{r})}. \tag{2.33}$$

$v_{\text{eff}}^{\text{KS}}(\boldsymbol{r})$ is equal to $v_s(\boldsymbol{r})$, the external potential in which the noninteracting electrons move. Therefore, for a given $v_{\text{eff}}^{\text{KS}}$, one can obtain the corresponding $\rho(\boldsymbol{r})$ by solving the N one-electron equations

$$\left[-\frac{1}{2}\nabla^2 + v_{\text{eff}}^{\text{KS}}(\boldsymbol{r}) \right] \psi_i = \varepsilon_i \psi_i, \tag{2.34}$$

with the condition

$$\rho(\boldsymbol{r}) = \sum_i^N \sum_s \left| \psi_i(\boldsymbol{r}, s) \right|^2. \tag{2.35}$$

Since $v_{\text{eff}}^{\text{KS}}$ depends on $\rho(\boldsymbol{r})$ through Eq. (2.33), Eq. (2.32) and Eq. (2.34) must be solved self-consistently. One constructs $v_{\text{eff}}^{\text{KS}}$ from Eq. (2.32) using a guessed $\rho(\boldsymbol{r})$, and then finds a new $\rho(\boldsymbol{r})$ from Eq. (2.34). The total energy can be computed directly from Eq. (2.30). Equations (2.32–2.35) constitute the KS equations.

2.2.3 More on the Kinetic-Energy Functional

The KS definition of $T_s[\rho]$ introduces a restriction in that it only holds for v-representable densities, meaning that a noninteracting reference system with the same electron density exists. However, one may also consider functionals defined on the larger domain of all N-representable densities. This can be achieved through the Levy constrained-search formulation[36,37] where $T_s[\rho]$ is defined by

$$T_s[\rho] = \min_{\Psi_{D\to\rho}} \left\langle \Psi_D \middle| \hat{T} \middle| \Psi_D \right\rangle, \tag{2.36}$$

which is followed by a second equality by expressing the kinetic-energy of the determinantal wave function Ψ_D by its N orbitals

$$T_s[\rho] = \min_{\sum |\psi_i^2| = \rho} \left[\sum_{i=1}^N \int \psi_i^*(\boldsymbol{r}) (-\frac{1}{2}\nabla^2) \psi_i(\boldsymbol{r}) d\boldsymbol{r} \right]. \tag{2.37}$$

The search is, hence, over all single-determinantal wave functions Ψ_D, which yield the given density ρ. $T_s[\rho]$ delivers the minimum of $\left\langle \Psi_D \middle| \hat{T} \middle| \Psi_D \right\rangle$, where $\hat{T}$ is the N-electron kinetic-energy operator.

In Eq. (2.37), one achieves a unique decomposition of a given N-representable density in terms of orbitals in the form of Eq. (2.35). This leads to the evaluation of the noninteracting kinetic-energy as the sum of the orbital kinetic-energies, in Eq. (2.29).

A somewhat different $T_s[\rho]$ can be defined by letting the search in Eq. (2.37) extend over all anti-symmetric wave functions Ψ rather than over determinants only. Such that,

$$\tilde{T}_s[\rho] = \min_{\Psi\to\rho} \left\langle \Psi \middle| \tilde{T} \middle| \Psi \right\rangle, \tag{2.38}$$

where

$$\tilde{T}_s[\rho] \leq T_s[\rho], \tag{2.39}$$

since the search in Eq. 2.38 is over a larger set than that in Eq. (2.37). However, for any non-interacting, v-representable density ρ that goes with a non-degenerate ground-state, it has been proven[36] that

$$\tilde{T}_s[\rho] = T_s[\rho]. \tag{2.40}$$

This can be shown as

$$\min_{\Psi\to\rho} \left\langle \Psi \middle| \tilde{T} \middle| \Psi \right\rangle = \min_{\Psi\to\rho} \left\langle \Psi \middle| \tilde{T} + \sum_{i=1}^{N} v_s(\boldsymbol{r}_i) \middle| \Psi \right\rangle - \int v_s(\boldsymbol{r})\rho(\boldsymbol{r})d\boldsymbol{r}. \tag{2.41}$$

Since ρ is non-interacting, v-representable, $v_s(\boldsymbol{r})$ in this equation can be chosen to be the potential for which $\rho(\boldsymbol{r})$ is the non-interacting, N-electron ground-state density. The minimum is achieved, consequently, at the corresponding determinantal wave function.

2.2.4 Exchange-Correlation Energy

The KS equations (2.32–2.35) provide an exact treatment of the noninteracting kinetic-energy $T_s[\rho]$, whereas the exchange-correlation functional $E_{\rm xc}[\rho]$ remains as an issue. An explicit form for $E_{\rm xc}$ is not known and needs to be approximated. There is not a definite strategy to search for an accurate exchange-correlation functional. Never the less, there are a number of available approximations, and the KS equations are open for improvement with each successive better approximation.

The simplest approximate exchange-correlation functional is the local-density approximation (LDA). The derivation of the LDA begins with dividing the space into numerous small cubes, each of side l and volume $V = l^3$, each containing some fixed number of electrons N. In these cells, a positive charge is uniformly spread to compensate the negative charge of the electrons. When one takes the limit $V \to \infty$ and $N \to \infty$, with $\rho = N/V$ is kept finite, one obtains the

model of uniform electronic gas. This model is used to obtain the LDA exchange-correlation energy

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[\rho] = \int \rho(\boldsymbol{r})\varepsilon_{\mathrm{xc}}(\rho)d\boldsymbol{r}, \tag{2.42}$$

where $\varepsilon_{\mathrm{xc}}(\rho)$ is the exchange-correlation energy per particle of a uniform electron gas of density ρ. Accurate values of $\varepsilon_{\mathrm{xc}}(\rho)$ are available, obtained from the quantum Monte Carlo calculations of Ceperley and Alder.[38] By interpolating these values, it is possible to obtain an analytic form for $\varepsilon_{\mathrm{xc}}(\rho)$. Application of the LDA to atoms and molecules assumes that the exchange-correlation for a non-uniform system can be obtained by applying uniform electron gas results to infinitesimall portions of the non-uniform distribution, each having $\rho(\boldsymbol{r})d\boldsymbol{r}$ electrons, and then summing over all space the individual contributions $\varepsilon_{\mathrm{xc}}(\rho)\rho(\boldsymbol{r})d\boldsymbol{r}$.

It is possible to improve this local-density approximation by considering the gradient of the density $\nabla\rho$. This leads to the generalized-gradient approximation (GGA):

$$E_{\mathrm{x}}^{\mathrm{GGA}}[\rho] = E_{\mathrm{xc}}^{\mathrm{LDA}}[\rho] \int d\boldsymbol{r}\rho^{4/3}F(s), \tag{2.43}$$

with

$$s = \frac{|\nabla\rho(\boldsymbol{r})|}{2k_F\rho}, \tag{2.44}$$

$$k_F = (3\pi^2\rho)^{1/3}, \tag{2.45}$$

where $F(s)$ is an enhancement factor depending on the density and its gradient. In most cases, the exchange-correlation functional is divided into an exchange part E_{x} and and a correlation part E_{c}, which are approximated separately. There are a number of widely used exchange-functionals, which include the functional of Becke[39] and of Perdew and Wang,[40] and for the correlation part, the functionals of Perdew,[41] and of Lee, Yang, and Parr.[42]

Another approach is to use the kinetic-energy density τ, which depends on the second derivative of the density, as a variable, in addition to the density and its gradient, which leads to the meta-GGA functionals.[43] There have also been developed functionals that in addition also depend on the occupied or even the virtual orbitals.[44–47]

It is noteworthy that the general approximations to the exchange-correlation (XC) functional fail to describe dispersion interactions,[48–51] which is a common contribution to the intermolecular interactions besides hydrogen bonding, electrostatic and exchange-repulsion. Intermolecular interactions play a pivotal role in structure formation of biomolecular systems such as protein folding, molecular

recognition and stacking of nucleobases. Therefore, there are numerous studies devoted to develop approximate DFT approaches that accurately treat them (for recent papers see Refs. 52–55).

2.2.5 Remarks on the KS Method

The KS method turns DFT into a practical tool for calculations. By introducing N orbitals, the KS equations handle $T_s[\rho]$ exactly, while the approximations are only introduced for $E_{\mathrm{xc}}[\rho]$. With each successive better approximation to $E_{\mathrm{xc}}[\rho]$, the KS equations can be improved, and would give the exact density and total energy if $E_{\mathrm{xc}}[\rho]$ was known precisely. Because compared to wave function based methods, the KS scheme provides a favourable balance between accuracy and efficiency, it has been used to treat numerous problems. This qualifies it as one of the most successful methods in quantum chemistry.

The price for exact treatment of $T_s[\rho]$ is that, there are N equations (orbitals), each a function of three variables, to solve as opposed to Hohenberg-Kohn theorem where only one equation of three variables for the total density derived from direct approximation on $T[\rho]$. Furthermore, the requirement that these N orbitals are orthogonal causes an eigenvalue problem, which calls for a diagonalization step that becomes particularly expensive for large systems due to the third power scaling with system size.

2.2.6 Orbital-Free DFT

In an orbital-free (OF) DFT scheme, no orbitals have to be calculated and no diagonalization step is needed, since the density is the sole variable to determine. This makes it a much more efficient scheme.

However, an explicit form of the the kinetic-energy functional in terms of the density only is not known. Therefore, approximations have to be introduced for the kinetic-energy in addition to the exchange-correlation functional. For the exchange-correlation energy, very accurate approximate functionals exist, that can be applied to a wide range of different systems. For example, the local-density approximation (LDA), which is the simplest possible approximation for both the exchange-correlation and kinetic energy as described in Sec. 2.2.4 is surprisingly accurate for the exchange-correlation energy. However, approximating the kinetic energy is much more difficult. For this reason, Kohn and Sham introduced their

exact treatment of the noninteracting kinetic energy, which made it for the first time possible to apply density-functional theory to molecules and turned DFT into an accurate and useful method.

It is more convenient to approximate the non-interacting kinetic-energy functional $T_s[\rho]$ than the true, interacting, kinetic-energy. The first reason for this is that is that the difference between T and T_s is included in the exchange-correlation energy, which, in most cases, is developed such that it includes this difference. If, in OF-DFT calculations, one approximates T instead of T_s, one has to find a new approximation for $E_{\rm xc}$. The second reason is that because the non-interacting kinetic-energy has a much simpler definition than the interacting kinetic-energy, simpler arguments can be used in the development of the approximate functionals.

There is a number of different approaches to adopt for developing approximate kinetic-energy functionals. A discussion of the available kinetic-energy functionals is given in Chapter 4. A brief overview as well as the discussion of the performance of these approximate kinetic-energy functionals can be also found in Ref.[SMB2] A more detailed overview is to be found in Ref. 56. The derivations and an extensive information about the theoretical background is given in Ref. 36. Below, a general information on the usage of the OF-DFT scheme is given.

In an OF-DFT scheme, the total energy functional in Eq. (2.30) can be minimised with respect to ρ, under the constraint that it integrates to the correct number of electrons and that the density is positive or zero at every point in space. In order to achieve this, the Lagrange multiplier μ is introduced, such that

$$0 = \frac{\delta}{\delta\rho}[E[\rho]] - \mu\Big(\int \rho(\boldsymbol{r})d\boldsymbol{r} - N\Big)] = \frac{\delta T_s[\rho]}{\delta\rho} + v_{\rm eff}^{\rm KS} - \mu, \tag{2.46}$$

with

$$\mu = \frac{\delta T_s[\rho]}{\delta\rho} + v_{\rm eff}^{\rm KS}[\rho], \tag{2.47}$$

where μ is identified with the orbital energy of the highest occupied KS orbital.[57] By solving this equation, one can obtain the ground-state electron density.

The density can be expressed in terms of a density orbital ψ,[58] i.e.

$$\rho(\boldsymbol{r}) = \psi(\boldsymbol{r})^2, \tag{2.48}$$

which is used as a new variational variable. This condition leads to ensuring the positivity of the electron density.

A part of the kinetic-energy functional can be expressed using the von Weizsäcker

approximation, as shown in Eq. (4.6), (details on this functional are given in Chapter 4.2),

$$T_s[\rho] = \frac{1}{8}\int \frac{|\nabla\rho(\boldsymbol{r})|^2}{\rho(\boldsymbol{r})d\boldsymbol{r}} + T_X[\rho] = \frac{1}{2}\int \nabla\psi(\boldsymbol{r})\nabla\psi(\boldsymbol{r})d\boldsymbol{r} + T_X[\rho], \qquad (2.49)$$

where T_X is the difference between T_s and the von Weizsäcker part $\frac{1}{8}\int \frac{|\nabla\rho(\boldsymbol{r})|^2}{\rho(\boldsymbol{r})d\boldsymbol{r}}$.

Incorporating Eq. (2.48) into Eq. (2.49), the functional derivative of the kinetic-energy functional can be written as

$$\frac{\delta T_s}{\delta\rho} = -\frac{1}{2}\frac{\nabla^2\psi}{\psi} + \frac{\delta T_X[\rho]}{\delta\rho}. \qquad (2.50)$$

Eq. (2.47) thus becomes

$$\left(-\frac{1}{2}\nabla^2 + \frac{\delta T_X[\rho]}{\delta\rho}\Big|_{\rho=\psi(\boldsymbol{r})^2} + v_{\text{eff}}[\psi^2](\boldsymbol{r}) \right)\psi(\boldsymbol{r}) = \mu\psi(\boldsymbol{r}). \qquad (2.51)$$

Even though this equation has a similar form as the KS equation, one has to solve it only for one orbital and not for a set of orthogonal orbitals. However, as long as the diagonalization step remains, such a scheme will also not be more efficient than KS-DFT calculations.

Chapter 3

Subsystem Density Functional Theory

In this Chapter, subsystem DFT that gives the foundation for the frozen-density embedding (FDE) scheme, is introduced. It begins with the essence of the theory, partitioning of the electron density. This is followed by a presentation of how subsystem DFT can be used to model large environments and how it provides an efficient alternative to conventional KS-DFT. Afterwards, it is explained how the kinetic-energy is approximated. Finally, a review of previous applications of subsystem DFT is given.

3.1 Partitioning of The Electron Density

In the FDE formalism,[59,60] the total electron density $\rho_{\mathrm{tot}}(\boldsymbol{r})$ is partitioned into the electron densities of two subsystems such that the two components $\rho_1(\boldsymbol{r})$ and $\rho_2(\boldsymbol{r})$ sum to the total density,

$$\rho_{\mathrm{tot}}(\boldsymbol{r}) = \rho_1(\boldsymbol{r}) + \rho_2(\boldsymbol{r}). \tag{3.1}$$

Provided that the subsystem densities integrate to an integer number of electrons, they are allowed to overlap. The nuclear charges are also partitioned accordingly, which together with the partitioning of the electron density defines two subsystems, subsystem 1 and subsystem 2.

Using this partitioning of the electron density, in the absence of any external

fields, the DFT total energy can be expressed as a functional of ρ_1 and ρ_2,

$$\begin{aligned} E[\rho_1,\rho_2] &= E_{\rm NN} + E_{\rm NA}[\rho_1,\rho_2] + E_{\rm Coulomb}[\rho_1,\rho_2] \\ &\quad + E_{\rm xc}[\rho_1+\rho_2] + T_s[\rho_1] + T_s[\rho_2] + T_s^{\rm nadd}[\rho_1,\rho_2], \end{aligned} \tag{3.2}$$

where $E_{\rm NN}$ is the nuclear repulsion energy, $E_{\rm NA}[\rho_1,\rho_2]$ is the nuclear attraction energy given by

$$\int (\rho_1(\boldsymbol{r}) + \rho_2(\boldsymbol{r}))(v_1^{\rm nuc}(\boldsymbol{r}) + v_2^{\rm nuc}(\boldsymbol{r}))d\boldsymbol{r}, \tag{3.3}$$

where $v_1^{\rm nuc}$ and $v_2^{\rm nuc}$ are the electrostatic potentials of the nuclei in subsystems 1 and 2, $E^{\rm Coulomb}[\rho_1,\rho_2]$ is the Coulomb energy defined as

$$\frac{1}{2}\int \frac{(\rho_1(\boldsymbol{r}) + \rho_2(\boldsymbol{r}))(\rho_1(\boldsymbol{r'}) + \rho_2(\boldsymbol{r'}))}{|\boldsymbol{r}-\boldsymbol{r'}|} d\boldsymbol{r}d\boldsymbol{r'}, \tag{3.4}$$

$E_{\rm xc}$ is the exchange-correlation energy functional, $T_s[\rho]$ is the kinetic-energy of the noninteracting reference system, and $T_s^{\rm nadd}$ is the nonadditive kinetic-energy defined as

$$T_s^{\rm nadd}[\rho_1,\rho_2] = T_s[\rho_1+\rho_2] - T_s[\rho_1] - T_s[\rho_2]. \tag{3.5}$$

Provided that the densities $\rho_1(\boldsymbol{r})$ and $\rho_2(\boldsymbol{r})$ are v_s-representable, canonical Kohn-Sham (KS) orbitals for the individual subsystems $\psi_i^{(n)}$ represent these densities, such that $\rho_n(\boldsymbol{r}) = 2\sum_{i=1}^{N_n/2}\psi_i^{(n)}(\boldsymbol{r})^*\psi_i^{(n)}(\boldsymbol{r})$ with $(n=1,2)$. Given this, it is possible to calculate the kinetic-energy of the corresponding non-interacting reference system as

$$T_s[\rho_n] = -2\sum_{i=1}^{N_n/2}\left\langle \psi_i^{(n)}\middle|\frac{\nabla^2}{2}\middle|\psi_i^{(n)}\right\rangle. \tag{3.6}$$

However, $T_s[\rho_1+\rho_2]$ cannot be calculated in this way because, as a consequence of the partitioning of the electron density, the representation of $\rho_{tot}(\boldsymbol{r})$ in the canonical KS orbitals is not available. This means that one has to calculate $T_s^{\rm nadd}[\rho_1,\rho_2]$ using an approximate kinetic-energy functional.

Similar to the kinetic-energy, the exchange-correlation energy in Eq. (3.2) can also be divided in the following way

$$E_{\rm xc}[\rho_1+\rho_2] = E_{\rm xc}[\rho_1] + E_{\rm xc}[\rho_2] + E_{\rm xc}^{\rm nadd}[\rho_1,\rho_2], \tag{3.7}$$

where $E_{\rm xc}^{\rm nadd}$ is defined as

$$E_{\rm xc}^{\rm nadd}[\rho_1,\rho_2] = E_{\rm xc}[\rho_1+\rho_2] - E_{\rm xc}[\rho_1] - E_{\rm xc}[\rho_2]. \tag{3.8}$$

This independent treatment of the subsystems enables one to use different approximations for the exchange-correlation functional for each subsystem. This leads to a subsystem specific approximation where the quality of the approximation used for the exchange-correlation functional depends on the subsystem.

3.2 The Embedding Potential

The minimisation of the total energy bifunctional in Eq. (3.2) with respect to $\rho_1(\boldsymbol{r})$ of the subsystem 1, while $\rho_2(\boldsymbol{r})$ of the subsystem 2 is kept frozen, enables one to determine the active electron density $\rho_1(\boldsymbol{r})$ for a given (frozen) $\rho_2(\boldsymbol{r})$. $\rho_1(\boldsymbol{r}) + \rho_2(\boldsymbol{r})$ equals the density that one would obtain from a conventional DFT calculation on the total system, with the condition that $\rho_1(\boldsymbol{r})$ of the active subsystem is v_s-representable[36] and positive everywhere.

Under the constraint that the number of electrons N_1 in subsystem 1 is conserved, when the total energy $E[\rho_1, \rho_2]$ is minimised with respect to ρ_1, one obtains the following condition

$$\begin{aligned} 0 &= \frac{\delta}{\delta\rho_1}\left[E[\rho_1,\rho_2] + \mu_1\left(\int \rho_1(\boldsymbol{r})d\boldsymbol{r} - N_1\right)\right] \\ &= v_1^{\mathrm{nuc}}(\boldsymbol{r}) + v_2^{nuc}(\boldsymbol{r}) + \int \frac{\rho_1(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r'}|}d\boldsymbol{r'} + \int \frac{\rho_2(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r'}|}d\boldsymbol{r'} \\ &\quad + \frac{\delta E_{\mathrm{xc}}[\rho_1+\rho_2]}{\delta\rho_1} + \frac{\delta T_s[\rho_1]}{\delta\rho_1} + \frac{\delta T_s^{\mathrm{nadd}}[\rho_1,\rho_2]}{\delta\rho_1} + \mu^{(1)}. \end{aligned} \tag{3.9}$$

For any v_s-representable[36] electron density $\rho_1(\boldsymbol{r})$ that is obtained from this minimisation, one can express $\rho_1(\boldsymbol{r}) = 2\sum_i^{N_1/2} |\psi_i^{(1)}(\boldsymbol{r})|^2$ where $\{\psi_i^{(1)}(\boldsymbol{r})\}$ are the KS orbitals of subsystem 1. The solution of the Kohn-Sham equations with constrained electron density (KSCED equations) determine these orbitals,

$$\Big[-\frac{\nabla^2}{2} + v_{\mathrm{eff}}^{\mathrm{KSCED}}[\rho_1,\rho_2](\boldsymbol{r})\Big]\psi_i^{(1)}(\boldsymbol{r}) = \varepsilon_i\psi_i^{(1)}(\boldsymbol{r}) \quad ; \quad i = 1,\ldots,N_1/2, \tag{3.10}$$

where $v_{\mathrm{eff}}^{\mathrm{KSCED}}$ is the effective potential of a noninteracting reference system, the electron density of which can be determined by minimising the noninteracting energy functional,

$$E_s[\rho_1] = T_s[\rho_1] + \int \rho_1(\boldsymbol{r}) v_{\mathrm{eff}}^{\mathrm{KSCED}}[\rho_1,\rho_2](\boldsymbol{r})d\boldsymbol{r}. \tag{3.11}$$

This minimisation is employed with respect to ρ_1, under the constraint that the correct number of electrons N_1 in subsystem 1 is conserved, leading to the condition

$$\begin{aligned} 0 &= \frac{\delta}{\delta\rho_1}\left[E_s[\rho_1] + \mu_1\left(\int \rho_1(\boldsymbol{r})d\boldsymbol{r} - N_1\right)\right] \\ &= \frac{\delta T_s[\rho_1]}{\delta\rho_1} + v_{\text{eff}}^{\text{KSCED}}[\rho_1,\rho_2](\boldsymbol{r}) + \mu_1. \end{aligned} \tag{3.12}$$

Using Eq. (3.9) and Eq. (3.12), one finds

$$v_{\text{eff}}^{\text{KSCED}}[\rho_1,\rho_2](\boldsymbol{r}) = v_{\text{eff}}^{\text{KS}}[\rho_{\text{tot}}](\boldsymbol{r}) + v_T[\rho_1,\rho_2](\boldsymbol{r}), \tag{3.13}$$

where $v_{\text{eff}}^{\text{KS}}[\rho_{\text{tot}}](\boldsymbol{r})$ is the KS effective potential of the total system. It can be decomposed into four pieces, namely

$$v_{\text{eff}}^{\text{KS}}[\rho_{\text{tot}}](\boldsymbol{r}) = v_1^{\text{nuc}}(\boldsymbol{r}) + v_2^{\text{nuc}}(\boldsymbol{r}) + \frac{\rho_{\text{tot}}(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r'}|}d\boldsymbol{r'} + \left.\frac{\delta E_{\text{xc}}[\rho]}{\delta\rho}\right|_{\rho=\rho_{\text{tot}}(\boldsymbol{r})}, \tag{3.14}$$

with the total nuclear potential, the Coulomb potential of the total electron density and the exchange-correlation potential of the total system.

$v_T[\rho_1,\rho_2]$ in Eq. (3.13) is a kinetic-energy component, which is defined as the functional derivative of the nonadditive kinetic-energy bifunctional,

$$v_T[\rho_1,\rho_2](\boldsymbol{r}) = \frac{\delta T_s^{\text{nadd}}[\rho_1,\rho_2]}{\delta\rho_1} = \left.\frac{\delta T_s[\rho]}{\delta\rho}\right|_{\rho=\rho_{\text{tot}}(\boldsymbol{r})} - \left.\frac{\delta T_s[\rho]}{\delta\rho}\right|_{\rho=\rho_1(\boldsymbol{r})}. \tag{3.15}$$

An approximate kinetic-energy functional should be used to model this kinetic-energy component v_T, for which the inequality $v_T[\rho_1,\rho_2] \neq v_T[\rho_2,\rho_1]$ holds, in contrast to the bifunctional of the nonadditive kinetic-energy in Eq. (3.5). This means that v_T is not symmetric with respect to the exchange of the two electron densities.

The terms of the KSCED effective potential in Eq. (3.14) can be separated such that those depending on the electron density and the positions of the nuclei of a given subsystem are taken together. This emphasises the effective embedding character of the FDE formalism, leading to the following expression

$$v_{\text{eff}}^{\text{KSCED}}[\rho_1,\rho_2](\boldsymbol{r}) = v_{\text{eff}}^{\text{KS}}[\rho_1](\boldsymbol{r}) + v_{\text{eff}}^{\text{emb}}[\rho_1,\rho_2](\boldsymbol{r}), \tag{3.16}$$

where $v_{\text{eff}}^{\text{KS}}[\rho_1](\boldsymbol{r})$ is the KS effective potential of the isolated subsystem 1 containing the nuclear potential, the Coulomb potential of the electrons, and the exchange-correlation potential,

$$v_{\text{eff}}^{\text{KS}}[\rho_1](\boldsymbol{r}) = v_1^{\text{nuc}}(\boldsymbol{r}) + \int \frac{\rho_1(\boldsymbol{r'})}{|\boldsymbol{r} - \boldsymbol{r'}|} d\boldsymbol{r'} + \left.\frac{\delta E_{\text{xc}}[\rho]}{\delta \rho}\right|_{\rho=\rho_1(\boldsymbol{r})} . \tag{3.17}$$

In Eq. (3.16), the effective embedding potential $v_{\text{eff}}^{\text{emb}}$ is given by the formula

$$v_{\text{eff}}^{\text{emb}}[\rho_1, \rho_2](\boldsymbol{r}) = v_2^{\text{nuc}}(\boldsymbol{r}) + \int \frac{\rho_2(\boldsymbol{r'})}{|\boldsymbol{r} - \boldsymbol{r'}|} d\boldsymbol{r'} + \frac{\delta E_{\text{xc}}[\rho_1, \rho_2]}{\delta \rho_1} + v_T[\rho_1, \rho_2](\boldsymbol{r}), \tag{3.18}$$

and it describes the interaction of subsystem 1 with the frozen density and nuclei of subsystem 2. $v_T[\rho_1, \rho_2](\boldsymbol{r})$ is the kinetic-energy component given in Eq. (3.15) and $\frac{\delta E_{\text{xc}}[\rho_1,\rho_2]}{\delta \rho_1}$ is the functional derivative of the nonadditive exchange-correlation energy in Eq. (3.8), and it reads

$$\frac{E_{\text{xc}}^{\text{nadd}}[\rho_1, \rho_2]}{\delta \rho_1} = \left.\frac{\delta E_{\text{xc}}[\rho]}{\delta \rho}\right|_{\rho=\rho_1+\rho_2} - \left.\frac{\delta E_{\text{xc}}[\rho]}{\delta \rho}\right|_{\rho=\rho_1} . \tag{3.19}$$

To summarise, by solving KS-like equations

$$\left[-\frac{\nabla^2}{2} + v_{\text{eff}}^{\text{KS}}[\rho_1](\boldsymbol{r}) + v_{\text{eff}}^{\text{emb}}[\rho_1, \rho_2](\boldsymbol{r}) \right] \psi_i^{(1)}(\boldsymbol{r}) = \varepsilon_i \psi_i^{(1)}(\boldsymbol{r}) \quad ; \quad i = 1, \ldots, N_1/2, \tag{3.20}$$

the density of the nonfrozen subsystem $\rho_1(\boldsymbol{r})$ can be determined for a given frozen density $\rho_2(\boldsymbol{r})$. As given in Eq. (3.18), $v_{\text{eff}}^{\text{emb}}$ represents the effect of the environment. Provided that ρ_1 is positive and v_s-representable, the solution of these equations yields the exact ground-state electron density of the total system.[60]

3.3 Modelling Large Environments

A small nonfrozen subsystem 1 embedded in a large environment (e.g. solvent) forms a typical system for FDE applications. It is particularly efficient to do molecular property calculations since these will only be performed on the small nonfrozen subsystem. However, it should be here emphasized that the applicability of FDE depends on the locality of the property. That is, only if $\frac{\partial \rho}{\partial \text{P}} = \frac{\partial \rho_1}{\partial \text{P}}$, where P is a given property and ρ_1 is the density of the nonfrozen subsystem.

When doing calculations on such large environments, obviously, obtaining the frozen density from a DFT calculation on the full system would be disadvantageous. Instead, one can also apply approximations in the construction of the environment.

There have been a number of studies on using approximate densities to model large environments. Among these, Wesolowski and Warshel[59,61] described a water environment using such an approximate density. Neugebauer *et. al*[62] investigated the electronic absorption of acetone in water using different approximate description of the frozen solvent environment. They found that using a superposition of gas phase densities for isolated water molecules gave results in good agreement with full DFT calculations. This strategy has been successfully used in more studies of solvent effects on molecular properties.[63,64] Including some solvent molecules in the nonfrozen subsystem[63] or to partially relax the frozen density with respect to the nonfrozen subsystem[65] were, however, necessary in some cases. For a good example where this technique fails see Ref. 66.

3.4 Subsystem Density Functional Theory

In practice, it is difficult to fulfil the condition that ρ_1 is v_s-representable at any point in space. This necessitates determining both the electron density in the nonfrozen and frozen subsystems, establishing the subsystem DFT formalism as it was initially proposed by Cortona.[67] Subsystem DFT provides a very efficient alternative to conventional KS-DFT.

In the subsystem DFT formalism, one minimises the total energy bifunctional in Eq. (3.2) not only with respect to the electron density ρ_1 in one of the subsystems while ρ_2 of the other subsystem is kept frozen, but instead applies an iterative scheme in which the electron densities of both subsystems are optimized. both subsystems. This pair of electron densities $\{\rho_1, \rho_2\}$ should be determined such that these densities minimise the total energy meaning that the energy is stationary with respect to variations in both densities,

$$0 = dE = \Big(\frac{\delta E}{\delta \rho_1}\Big)\delta\rho_1 + \Big(\frac{\delta E}{\delta \rho_2}\Big)\delta\rho_2 \quad \forall \delta\rho_1, \delta\rho_2, \tag{3.21}$$

where the functional derivatives with respect to ρ_1 and ρ_2 are partial derivatives, while the other density is kept frozen. The variations $\delta\rho_1$ and $\delta\rho_2$ have to conserve the number of electrons N_1 and N_2 in the individual subsystems, respectively.

The total energy is then stationary with respect to variations of $\rho_{\mathrm{tot}} = \rho_1 + \rho_2$ because of the fact that it is stationary with respect to variations of both ρ_1 and ρ_2. However, the partitioning of the total density is not unique. This means that, even if the number of electrons are fixed, there will be many partitionings of the total electron density minimising the total energy, since the total energy only depends on the total electron density ρ_{tot}.

Employing the stationary condition given in Eq. (3.21) yields for the subsystem densities

$$0 = \frac{\delta}{\delta\rho_1}\left[E[\rho_1,\rho_2] + \mu_1\left(\int \rho_1(\boldsymbol{r})d\boldsymbol{r} - N_1\right)\right] \tag{3.22}$$

$$0 = \frac{\delta}{\delta\rho_2}\left[E[\rho_1,\rho_2] + \mu_2\left(\int \rho_2(\boldsymbol{r})d\boldsymbol{r} - N_2\right)\right] \tag{3.23}$$

This leads to two coupled sets of KSCED equations,

$$\left[-\frac{\nabla^2}{2} + v_{\mathrm{eff}}^{\mathrm{KS}}[\rho_1](\boldsymbol{r}) + v_{\mathrm{eff}}^{\mathrm{emb}}[\rho_1,\rho_2](\boldsymbol{r})\right]\psi_i^{(1)}(\boldsymbol{r}) = \varepsilon_i^{(1)}\psi_i^{(1)}(\boldsymbol{r}), \quad i = 1,\ldots,N_1/2, \tag{3.24}$$

$$\left[-\frac{\nabla^2}{2} + v_{\mathrm{eff}}^{\mathrm{KS}}[\rho_2](\boldsymbol{r}) + v_{\mathrm{eff}}^{\mathrm{emb}}[\rho_2,\rho_1](\boldsymbol{r})\right]\psi_i^{(2)}(\boldsymbol{r}) = \varepsilon_i^{(2)}(\boldsymbol{r})\psi_i^{(2)}(\boldsymbol{r}), \quad i = 1,\ldots,N_2/2, \tag{3.25}$$

where $v_{\mathrm{eff}}^{\mathrm{KS}}[\rho_1]$ and $v_{\mathrm{eff}}^{\mathrm{KS}}[\rho_2]$ are the KS potentials of the isolated subsystem 1 and 2, respectively. $v_{\mathrm{eff}}^{\mathrm{emb}}$ is the effective embedding potential defined in Eq. (3.18).

These two equations are coupled and have to be solved iteratively since $v_{\mathrm{eff}}^{\mathrm{emb}}$ depends on the electron densities of both subsystems. This can be achieved by employing the so-called "freeze-and-thaw" cycles,[68] where the roles of the frozen and nonfrozen subsystem are interchanged until the convergence is reached. However, these "freeze-and-thaw" cycles are not needed if the condition that $\rho_1 = \rho_{\mathrm{tot}} - \rho_2$ is fulfilled and that the validity of an exact functional is in discussion[58,69]

It is trivial to extend the subsystem DFT scheme to an arbitrary number of subsystems where the total electron density of M number of subsystems is partitioned as

$$\rho_{\mathrm{tot}}(\boldsymbol{r}) = \sum_{i=1}^{M} \rho_i(\boldsymbol{r}). \tag{3.26}$$

In this case, a set of M coupled KSCED equations is obtained, in which the frozen density in the effective embedding potential is replaced by the sum of the densities of all frozen subsystems.[67,70]

Details of a flexible implementation of the FDE scheme allowing the subsystem DFT treatment with many subsystems can be found in Ref. 27.

3.5 Approximating the Nonadditive Kinetic-Energy

The choice of the approximation which is used for the nonadditive kinetic-energy component that both the total energy bifunctional and the effective embedding potential contain influences significantly the performance of the FDE scheme.

The nonadditive kinetic-energy is approximated in the following way

$$\tilde{T}_s^{\mathrm{nadd}}[\rho_1,\rho_2] = \tilde{T}_s[\rho_1,\rho_2] - \tilde{T}_s[\rho_1] - \tilde{T}_s[\rho_2], \tag{3.27}$$

accordingly, the kinetic energy component v_T of the embedding potential is approximated as

$$\left.\frac{\tilde{T}_s[\rho]}{\delta\rho}\right|_{\rho=\rho_{\mathrm{tot}}(\boldsymbol{r})} - \left.\frac{\delta\tilde{T}_s[\rho]}{\delta\rho}\right|_{\rho=\rho_1(\boldsymbol{r})} \tag{3.28}$$

where tilde labels the approximate quantities, and $\tilde{T}_s[\rho]$ refers to an approximate kinetic-energy functional.

It is important to note that in the subsystem DFT scheme, the approximate kinetic-energy functional describes only a relatively small part of the kinetic-energy, while the kinetic-energies corresponding to two components of the density ρ_1 and ρ_2 are treated exactly using the corresponding KS orbitals. That is, the total noninteracting kinetic-energy can be written as

$$\begin{aligned} T_s[\{\psi_i^{(1)}\},\{\psi_i^{(1)}\}] \;=\; & \sum_{i=1}^{N}\int \psi_i^{(1)}(\boldsymbol{r})^*\left(\frac{-\nabla^2}{2}\right)\psi_i^{(1)}(\boldsymbol{r})d\boldsymbol{r} \\ & + \sum_{i=1}^{N}\int \psi_i^{(2)}(\boldsymbol{r})^*\left(\frac{-\nabla^2}{2}\right)\psi_i^{(2)}(\boldsymbol{r})d\boldsymbol{r} \\ & + \tilde{T}_s^{\mathrm{nadd}}[\rho_1,\rho_2]. \end{aligned} \tag{3.29}$$

This distinguishes the requirements on the approximate functionals for the use in subsystem DFT from those in orbital-free DFT. This means that, functionals that

perform poorly in orbital-free DFT can still provide good results in subsystem DFT.[71–73] Provided that the exact kinetic-energy functional $T_s[\rho]$ is used, for a given approximation to the exchange-correlation energy $E_{\rm xc}[\rho]$, subsystem DFT should reproduce the supermolecular KS results.

As mentioned in the earlier chapter, the overview of different approximate kinetic-energy functionals can be found in Chapter 4.

3.5.1 Review of Applications of Subsystem DFT

The roots of subsystem formulation of DFT can be traced back to 1986[74] but only in 1991 and later in 1991 the subsystem formulation of DFT was proposed and which provided the foundation of the FDE scheme, which was proposed in 1993. In this chapter, an overview of the applications of subsystem DFT will be given dating from 1991 until the present time. A more complete overview dating until 2006 can be found in Ref.[60,72,73]

Studies on interaction energies Wesolowski and Warshel investigated the solvation of a lithium ion in water and the solvation free energy of liquid water and methane,[21,61] where the frozen solvent environments were constructed by using a superposition of the densities of isolated water or methane molecules as described in Chapter 3.3.

Furthermore, Wesolowski and co-workers investigated the accuracy of different kinetic-energy functionals for hydrogen-bound complexes by comparing interaction energies obtained from "freeze-and-thaw" calculations to those obtained from conventional KS-DFT calculations.[68,71–73] The studies on different sets of hydrogen-bound[75] and non-covalently bound intermolecular complexes[76] for interaction energies, as well as equilibrium structures[77] followed up this work. However, these later studies showed that error cancellations might hinder the problems and inaccuracies in the nonadditive kinetic-energy functionals used. A similar error cancellation was also observed in studies on van der Waals complexes,[78–80] where FDE "freeze-and-thaw" calculations of the interaction energies outperformed conventional KS-DFT calculations. In a more recent work, strong interaction energies such as in ligand-metal bonds[SMB2] were studied. The followup up of this study where the dispersion correction was employed is given in Chapter 6 (Ref.[SMB3]).

Studies on embedding potentials In a different type of study on van der Waals complexes, the possible problems in the kinetic-energy functionals were identified,[81] by comparing the dipole moments obtained in "freeze-and-thaw" FDE calculations with those of conventional KS-DFT. Kiewisch *et al.*[82] compared the the electron densities for a set of hydrogen-bonded complexes obtained from subsystem DFT calculations, where different approximations to the kinetic-energy were employed, with those of conventional KS-DFT. They found out that subsystem DFT is able to reproduce the characteristics of the density in the bonding region even for the strong hydrogen-bonden system F–H–F^-.[82] However, they also underlined the issues concerning the currently available approximations to the kinetic-energy functional.

The identified problems in the kinetic-energy component of the embedding potential were later in more detail investigated.[SMB4] In this study, a position-dependent correction has been proposed which ensures the correct behaviour for this kinetic-energy component at the long distance limit. Fux *et al.*[83] employed this correction to analyse the electron density distributions from FDE calculations on subsystems connected by coordination bonds. They found that this correction leads to an improved description of covalent bonds in transition metal complexes, even though the resulting density is still not very accurate. In a different study,[84] a non-decomposable approximation for the non-additive kinetic-energy component has been constructed which also enforces the exact limit near nuclei in the environment where the density of the active subsystem approaches to zero.

In a different study, the accuracy of a representative set of currently available approximate kinetic-energy functionals used within the subsystem DFT scheme have been assessed for NgAuF (Ng = Ar, Kr, Xe) molecules, which were partitioned into a Ng and a AuF subsystem.[SMB5] The effective-embedding potentials and resulting electron density distributions have been analysed. The results of this study showed that none of the tested approximate kinetic-energy functionals performs well enough to describe the weak covalent bond between the noble gas and gold adequately, contributes to the growing evidence that the current procedure to obtain approximate kinetic-energy functionals by reparametrizing functionals obtained via the conjointness hypothesis of Lee, Lee, and Parr[85] is insufficient to treat metal-ligand interactions with covalent character.

In order to obtain improved approximations to the nonadditive kinetic-energy functional, so that subsystem DFT can directly treat covalently-bonded subsystems, Fux *et al.*[86] have implemented a method for the numerical calculation of accurate references for the kinetic-energy component of the embedding potential. They analysed these accurate potentials and compared them to those obtained

from popular kinetic-energy density functionals, for a set of model systems consisting of subsystems connected by hydrogen bonds of various strengths, a coordination bond a covalent bond.

This above mentioned method has also been used on a different system, a Li_2 molecule, partitioned into a Li^+ and a Li^- subsystem, where the Li^+ molecule was treated as the frozen subsystem at the initial embedding step. This way a minimum negative density, defined as the negative part of the difference between the total density of the supermolecule and that of the frozen subsystem, has been ensured. In this study, the effect of "freeze-and-thaw" cycles, the necessity of which was investigated earlier,[69] on the reconstructed potentials and electron densities have been assessed. LiH molecule has also been studied, partitioned into a Li^+ and H^- subsystem. This system is particularly interesting since it is possible to calculate the exact effective embedding potential for the H^- subsystem, which then can be compared with the reconstructed one. The results of this study are presented in Chapter 5.

Jacob and co-workers[87] have presented an extension of the FDE scheme that can be applied to subsystems connected by covalent bonds which can be employed for quantum chemical calculations of proteins by treating each constituting amino acid as a separate subsystem. They assessed the accuracy of this extended FDE scheme by performing calculations for several dipeptides and for the protein ubiquitin.

Studies on molecular properties of embedded systems The Subsystem DFT scheme has been also employed to determine various molecular properties of molecules in different environments. Wesolowski investigated the electron spin resonance (ESR) hyperfine coupling constants (hfccs) of Mg^+ in Ne and Ar matrices, where the matrix environment was represented by a small cluster of rare gas atoms.[88] Neugebauer *et al.* also investigated ESR hfccs, who studied H_2NO in water.[64] In both cases, the calculated hfccs were in good agreement with the experimental values.

Neugebauer and co-workers have looked at solvent effects on different molecular properties where they approximated the solvent environment as a sum of the densities of isolated solvent molecules. They investigated solvatochromic shifts for acetone in water[62] and for the organic dye aminocoumarin C151 in n-hexane and in water.[63] In a related study, a systematic comparison of the FDE scheme to QM/MM methods for the description of solvent effects of different molecular properties has been presented.[65]

Bulo *et al.*[89] presented a subsystem DFT study of solvent effects on nuclear magnetic shielding parameters, where they studied nitrogen shift of acetonitrile immersed in a selected set of solvents, namely water, chloroform, and cyclohexane. Comparison of their FDE results showed close agreement with those of conventional supermolecular DFT and experiments for small-solvent and large-solvent clusters; respectively.

There are also several different applications where the time-dependent DFT (TDDFT) extension of FDE,[90] initially benchmarked for DNA base pairs,[91] was employed in several different applications. Among those are the above mentioned studies of solvatochromic shifts,[62,63,65] and for the calculation of polarisabilities.[65] Neugebauer *et al.*[92] used it to describe induced circular dichroism (CD) in host-guest systems. However, since FDE cannot describe couplings between excitations in the different subsystems, they found cavity problems for a system consisting of phenole in a cyclodextrin. These problems were later solved by further improvements of the subsystem TDDFT formalism by Neugebauer in a later work.[93]

Carter and co-workers[94,95] have extended the FDE formalism to a wave-function theory (WFT) -in-DFT embedding scheme. They applied this WFT-inDFT embedding scheme to study the adsorption of CO on a Cu(111) surface.[94,95] They also described localised electronic excitations in a CO molecule adsorbed on a Pd(111) surface[96,97] and investigated the adsorption of a cobalt atom on a Cu(111) surface.[98]

Gomes *et al.*[66] developed a slightly different scheme, where the effective embedding potential is obtained from a standard FDE calculation and then subsequently is used in a wave-function calculation on the active subsystem. This scheme requires only a single wave-function calculation since the embedding potential is constructed using the DFT density for the active subsystem. In this case, however, this DFT density of the active subsystem must be accurate enough.

Chapter 4

The weak covalent bond in NgAuF (Ng = Ar, Kr, Xe): A challenge for subsystem density functional theory [1]

Abstract

We have assessed the accuracy of a representative set of currently available approximate kinetic-energy functionals used within the frozen-density embedding scheme for the NgAuF (Ng=Ar, Kr, Xe) molecules, which we partitioned into a Ng and a AuF subsystem. Although it is weak, there is a covalent interaction between these subsystems which represents a challenge for this subsystem density functional theory approach. We analyzed the effective-embedding potentials and resulting electron density distributions and provide a quantitative analysis of the latter from dipole moment differences and root-mean-square errors in the density with respect to the supermolecular Kohn-Sham density functional theory reference calculation. Our results lead to the conclusion that none of the tested approximate kinetic-energy functionals performs well enough to describe the bond between the noble gas and gold adequately. This observation contributes to the growing evidence that the current procedure to obtain approximate kinetic-energy functionals by reparametrizing functionals obtained via the "conjointness" hy-

[1] The contents of this chapter have been published as S. M. Beyhan, A. W. Götz, C. R. Jacob, and L. Visscher, *J. Chem. Phys.* **132**, 044114 (2010)

pothesis of Lee, Lee and Parr [H. Lee, C. Lee, and R. Parr, Phys. Rev. A **44**, 768 (1991)] is insufficient to treat metal-ligand interactions with covalent character.

4.1 Introduction

The increasing interest in application of quantum chemical methods in the study of biological systems has lead to widespread use of subsystem methods that focus attention on a particular region of interest, treating the environment at a lower level of theory.[3–6] An attractive realization of such a subsystem approach is the so-called frozen–density embedding (FDE) method within density-functional theory (DFT) as proposed by Wesolowski and Warshel.[59] In this method the electron density of the environment is included in the calculation by means of an effective embedding potential that describes the difference between the full and the active system. This method can be used to calculate molecular properties of solvated systems[62,65,89] and, in its generalization to time-dependent DFT,[90,93] to describe local electronic excitations and couplings between such excitations.[63,66,99,100] The method can also be used to compute interaction energies of weakly interacting systems[101–103SMB2] as well as strong interaction energies as in ligand-metal bonds[SMB2] or to optimize structures of such complexes[77] and has even been explored in molecular dynamics simulations.[70]

The foundation of FDE is a subsystem formulation of DFT.[67] In most cases the density is partitioned into two subdensities that each correspond to an integer number of electrons, but a three-partitioning scheme that uses capping atoms to enable fractionation of strongly interacting subsystems, has also been developed.[87] In FDE, the density $\rho^{(1)}(\boldsymbol{r})$ of an active fragment is determined in the presence of an effective embedding potential due to the frozen electron density $\rho^{(2)}(\boldsymbol{r})$ of the environment. Provided that an initial guess for the environment density $\rho^{(2)}(\boldsymbol{r})$ is available, Kohn-Sham-like one-electron equations can be obtained for the determination of this active density $\rho^{(1)}(\boldsymbol{r})$ from the minimization of the energy functional $E = E[\rho^{(1)}, \rho^{(2)}]$ with respect to $\rho^{(1)}(\boldsymbol{r})$, while keeping $\rho^{(2)}(\boldsymbol{r})$ frozen.[59,60] To stress the difference with regular Kohn-Sham approaches these one-electron equations can be called the Kohn-Sham (KS) equations with constrained electron density (KSCED). For doubly occupied orbitals $\phi_i^{(1)}(\boldsymbol{r})$ of subsystem 1 they read

$$\Big[-\frac{\nabla^2}{2} + v_{\text{eff}}^{\text{KSCED}}[\rho^{(1)}, \rho^{(2)}](\boldsymbol{r}) \Big] \phi_i^{(1)}(\boldsymbol{r}) = \epsilon_i \phi_i^{(1)}(\boldsymbol{r}), \quad i = 1, \ldots, \frac{N^{(1)}}{2}. \tag{4.1}$$

The effective potential in these equations is given by

$$v_{\text{eff}}^{\text{KSCED}}[\rho^{(1)}, \rho^{(2)}](\boldsymbol{r}) = v_{\text{eff}}^{\text{KS}}[\rho^{(1)}](\boldsymbol{r}) + v_{\text{eff}}^{\text{emb}}[\rho^{(1)}, \rho^{(2)}](\boldsymbol{r}), \tag{4.2}$$

where $v_{\text{eff}}^{\text{KS}}[\rho^{(1)}](\boldsymbol{r})$ is the KS effective potential of the isolated subsystem 1 containing the usual terms of the nuclear potential, the Coulomb potential of the elec-

trons, and the exchange-correlation potential,

$$v_{\mathrm{eff}}^{\mathrm{KS}}[\rho^{(1)}](\boldsymbol{r}) = v_{\mathrm{nuc}}^{(1)}(\boldsymbol{r}) + \int \frac{\rho^{(1)}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r}' + \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta\rho}\Big|_{\rho=\rho^{(1)}(\boldsymbol{r})}. \tag{4.3}$$

The effect of subsystem 2 is represented by the effective-embedding potential $v_{\mathrm{eff}}^{\mathrm{emb}}[\rho^{(1)}, \rho^{(2)}](\boldsymbol{r})$ that reads

$$\begin{aligned} v_{\mathrm{eff}}^{\mathrm{emb}}[\rho^{(1)}, \rho^{(2)}](\boldsymbol{r}) &= v_{\mathrm{nuc}}^{(2)}(\boldsymbol{r}) + \int \frac{\rho^{(2)}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r}' \\ &\quad + \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta\rho}\Big|_{\rho=\rho^{(\mathrm{tot})}(\boldsymbol{r})} - \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta\rho}\Big|_{\rho=\rho^{(1)}(\boldsymbol{r})} \\ &\quad + \frac{\delta T_{\mathrm{s}}[\rho]}{\delta\rho}\Big|_{\rho=\rho^{(\mathrm{tot})}(\boldsymbol{r})} - \frac{\delta T_{\mathrm{s}}[\rho]}{\delta\rho}\Big|_{\rho=\rho^{(1)}(\boldsymbol{r})}, \end{aligned} \tag{4.4}$$

where $v_{\mathrm{nuc}}^{(2)}(\boldsymbol{r})$ denotes the external potential due to the nuclei of system 2, $\rho^{(\mathrm{tot})}(\boldsymbol{r}) = \rho^{(1)}(\boldsymbol{r}) + \rho^{(2)}(\boldsymbol{r})$ is the electron density of the whole system, $T_{\mathrm{s}}[\rho]$ is the functional for the kinetic energy of the non-interacting reference system defined in the KS theory, and $E_{\mathrm{xc}}[\rho]$ is the functional for the exchange-correlation energy.

The formalism assumes that, for a given $\rho^{(2)}(\boldsymbol{r})$, the active density $\rho^{(1)}(\boldsymbol{r}) = \rho^{(\mathrm{tot})}(\boldsymbol{r}) - \rho^{(2)}(\boldsymbol{r})$ is non-negative everywhere in space and non-interacting v_{s}-representable.[60,82] In that case the supermolecular KS-DFT results for a given approximate functional for $E_{\mathrm{xc}}[\rho]$ should be reproduced by the subsystem calculation, provided that the exact kinetic-energy functional $T_s[\rho]$ is used. In practice the first condition is difficult to fulfill exactly with a simple trial density, making it necessary to introduce the so-called "freeze-and-thaw" (FT) procedure[68] in which both densities are adjusted in an iterative fashion. More importantly, the exact $T_{\mathrm{s}}[\rho]$ is unknown and one has to resort to an approximant for the non-additive kinetic-energy and the kinetic-energy component of the embedding potential. Wesolowski and coworkers have examined such approximate kinetic-energy functionals in the FDE scheme[71–73,78,81,104] and have shown that with generalized-gradient approximation (GGA) functionals, accurate results are obtained for a variety of weakly interacting systems.[78,81,104] A similar conclusion was reached by Kiewisch *et al.*[82] who studied also the strongly hydrogen-bonded system F–H–F^-.[82] Recently, however, Fux *et al.*[83] analyzed the electron density distributions from FDE calculations on subsystems connected by coordination bonds. They show that FDE, with a GGA approximation to the kinetic-energy

component of the embedding potential, fails for compounds with strong covalent bonding contributions.

In an earlier study,[SMB4] an exact form for the non-additive kinetic-energy component of the effective embedding potential $v_{\text{eff}}^{\text{emb}}[\rho^{(1)}, \rho^{(2)}](\boldsymbol{r})$ at the long-distance limit has been derived and a position-dependent correction has been proposed which ensures the correct behavior for this kinetic-energy component in this limit. This correction has been shown to lead to an improved description of covalent bonds in transition metal complexes, even though the resulting density is still not very accurate.[83] More recently, an approximation for the non-additive kinetic-energy component has been constructed which enforces the exact limit near nuclei in the environment.[84]

Our goal is to analyze the behavior of the currently available GGAs that underly both the older and the newer generations of the non-additive kinetic-energy functionals. These functionals are constructed following successful exchange functional forms, in line with the conjointness hypothesis of Lee, Lee and Parr (LLP).[85] In our analysis, we also include the PBEn kinetic-energy functionals developed by Karasiev *et al.*[105] that have only recently been tested in the context of FDE.[SMB2]

The present study was designed to assess the NgAuF (Ng=Ar, Kr, Xe) molecules in which the bond between the noble gas and AuF has considerably covalent character[106,107] and for which we may tune the interaction strength by changing the coordinating noble gas atom. Another advantage of this type of molecules is their linearity, making it easy to visualize the embedding potentials and deformation densities along the bond axis.

In Section 2, we briefly discuss the available approximate kinetic-energy functionals and the conjointness conjecture. Section 3 contains details on the computational methods applied in this work. Section 4 is devoted to discussion of the results. In Section 4.1, we examine the differences between the embedding potentials calculated using different approximations for the non-additive kinetic-energy functional. In Section 4.2, we consider the induced dipole moments of the Ng (Ng=Ar, Kr, Xe) atoms that are obtained without relaxing the density of the AuF unit. In Section 4.3, we allow for this relaxation using the FT procedure and in Section 4.4 we quantify differences of the obtained total density with respect to the supermolecular reference density. Finally, concluding remarks are given in Section 5.

4.2 Approximate Kinetic-Energy Functionals

The simplest approach to describe the kinetic energy in terms of a density functional is the Thomas-Fermi (TF) model.[108–110] The TF model has well-known defects in the description of the total kinetic energy of molecules[111–114] but yields a reasonable approximation to the repulsive part of the embedding potential to be of use as a starting point for the FDE approach. The TF kinetic-energy functional is given by

$$T_{\mathrm{TF}}[\rho] \;=\; C_{\mathrm{TF}} \int \rho^{5/3}(\boldsymbol{r}) d\boldsymbol{r}, \qquad C_{\mathrm{TF}} = \frac{3}{10}(3\pi^2)^{2/3} \approx 2.871 \tag{4.5}$$

where C_{TF} is the TF constant. This local density approximation (LDA) of the kinetic energy is exact for the uniform-electron gas.

Another simple model was developed by von Weizsäcker (vW),[115] based on another exact limit: the kinetic-energy density functional for 2-electron systems. This functional is defined by the equation

$$T_{\mathrm{vW}}[\rho] \;=\; \frac{1}{8} \int \frac{|\nabla \rho(\boldsymbol{r})|^2}{\rho(\boldsymbol{r})} d\boldsymbol{r}. \tag{4.6}$$

The vW functional should be applicable in the outer regions of a molecule and close to the nuclei, where the density can be considered to be due to a single orbital. Since it gives zero kinetic energy in the uniform-electron gas limit it can be included as a correction to the TF kinetic energy. This gives the Thomas-Fermi-von Weizsäcker model (TFW)

$$T_{\mathrm{TFW}}[\rho] \;=\; T_{\mathrm{TF}}[\rho] + \lambda T_{\mathrm{W}}[\rho] \tag{4.7}$$

where λ is a parameter with $0 \leq \lambda \leq 1$. The standard value $\lambda = 1/9$ follows from a second-order gradient expansion of the kinetic energy. The TFW model has been shown to be able to improve the overall accuracy relative to the individual functionals, but the errors in the total kinetic energy are still too large to be used in quantitative calculations.

Further sophistications can be introduced to model deviations from the uniform-electron gas limit by including the gradients of the density in GGA functionals. Modern developments thereby usually follow the idea of Lee, Lee and Parr (LLP) who conjectured the concept of "conjointness"[85] of the scale invariant part of the kinetic and exchange energy expressions. This idea has been widely adopted, with examples given in Refs.[116–120] According to the LLP conjointness concept,

kinetic-energy functionals are written in the form of a GGA,[121]

$$T_{\mathrm{s}}^{\mathrm{GGA}}[\rho] = C_{\mathrm{TF}} \int \rho^{5/3}(\boldsymbol{r}) F_{\mathrm{t}}(s(\boldsymbol{r})) d\boldsymbol{r}, \tag{4.8}$$

where the dimensionless function $F_{\mathrm{t}}(s)$ is called the "enhancement factor" and the reduced density gradient is denoted by $s \equiv |\nabla\rho|/(2\rho k_{\mathrm{F}})$, with $k_{\mathrm{F}} = (3\pi^2\rho)^{1/3}$. For the enhancement factor one then uses the same functional form as for approximate exchange-functionals, i.e.,

$$F_{\mathrm{t}}(s(\boldsymbol{r})) \approx F_{\mathrm{x}}(s(\boldsymbol{r})). \tag{4.9}$$

In practical orbital-free DFT (and FDE) calculations the parameters in the enhancement factor are often refitted to yield an improved description of the (non-additive) kinetic energy and its functional derivative, the kinetic contribution to the (embedding) potential. In this work we will compare the behavior of five different GGA functionals based on this conjointness concept.

The PW91K GGA kinetic-energy functional[72,116] has the same functional form for the enhancement factor $F(s)$ as the exchange functional of Perdew and Wang (PW91),[40] and was parametrized for the kinetic-energy by Lembarki and Chermette[116] as:

$$F_{\mathrm{t}}^{\mathrm{PW91K}}(s) = \frac{1 + A_1 s \cdot \sinh^{-1}(\mathrm{A}s) + (\mathrm{A}_2 - \mathrm{A}_3 \mathrm{e}^{-\mathrm{A}_4 s^2}) s^2}{1 + A_1 s \cdot \sinh^{-1}(\mathrm{A}s) + \mathrm{B}_1 s^4}, \tag{4.10}$$

with A_1 = 0.093907, A_2 = 0.26608, A_3 = 0.0809615, A_4 = 100.00, A = 76.320, and B_1 = 0.57767 $\cdot$ 10^{-4}.

This functional form of the kinetic energy is useful for embedding purposes because $F_{\mathrm{t}}^{\mathrm{PW91K}}(s)$ smoothly approaches zero as s becomes large.[60] This makes the kinetic-energy contribution to the embedding potential positive in regions of low density of the active subsystem since the last term of Eq. 4.4 remains small.

The PW91 enhancement factor is, however, rather complicated. For exchange-correlation functionals it is often replaced by the simple function introduced for E_{x} by Becke and used by Perdew, Burke, and Ernzerhof in their PBE exchange-correlation functional.[122] This enhancement function $F_t(s)$

$$F_{\mathrm{t}}^{\mathrm{PBE}}(s) = 1 + \frac{C_1 s^2}{1 + a_1 s^2}, \tag{4.11}$$

has also been reparametrized for use in kinetic energy functionals. Tran and Wesolowski fitted the parameters to reproduce the exact kinetic energy of the He

and Xe atoms and developed the TW02 functional that has C_1 = 0.2319 and a_1 = 0.2748. Karasiev, Trickey and Harris[105] chose to reproduce the KS forces of a training set of silicon oxide molecules, irrespective of the resulting total energy. Since the total kinetic energy is not of interest in FDE calculations (as the bulk of the energy is given by the KS expression), and because the functional derivative appearing in the force expression also appears in the expression for the embedding potential, this ansatz may also be interesting in the present context. The resulting parameters (C_1 = 8.7575 and a_1 = 1.0706) are substantially larger than the TW02 parametrization making this PBE2 functional deviate considerably from the TF starting point. It is thus interesting to see what such a larger departure from the TF starting point may give, although we should keep in mind that the training set (silicon oxide bonds) used by Karasiev and coworkers is quite different from the weak covalent bonding that we are aiming to treat.

The same authors[105] also developed three- and four-parameter enhancements factors $F(s)$ (dubbed PBE3 and PBE4, respectively, by them) based on the expressions introduced by Adamo and Barone:[123]

$$F_{\mathrm{t}}^{\mathrm{PBEn}}(s) = 1 + \sum_{i=1}^{n-1} C_i^{(n)} \left[\frac{s^2}{1 + a_1^{(n)} s^2} \right]^i , \tag{4.12}$$

with $C_1^{(3)} = -3.7425$, $a_1^{(3)}$ = 4.1355 and $C_2^{(3)}$ = 50.258, and $C_1^{(4)} = -7.2333$, $a_1^{(4)}$ = 1.7107, $C_2^{(4)}$ = 61.645 and $C_3^{(4)} = -93.683$.

In contrast to the functions discussed earlier, the PBE3 and PBE4 enhancement factors may attain values smaller than one for small values of s, thus reducing the TF kinetic energy density rather than enhancing it. This difference between the enhancement factors is displayed in Figure 1. Apart from the small s behavior of PBE3 and PBE4 one may also note the similarity between the PW91K and TW02 enhancement functions.

4.3 Computational Details

All calculations were performed using the FDE[59] implementation[63,124] in the Amsterdam Density Functional (ADF) package.[125,126] The PBE exchange-correlation functional,[122] the zeroth order regular approximation (ZORA)[127,128] and the TZ2P basis set from the ADF basis set library[126] were employed throughout this work. FDE calculations were performed using both the default basis set expansion (denoted as FDE(m)), in which only the basis functions of the active subsystem are

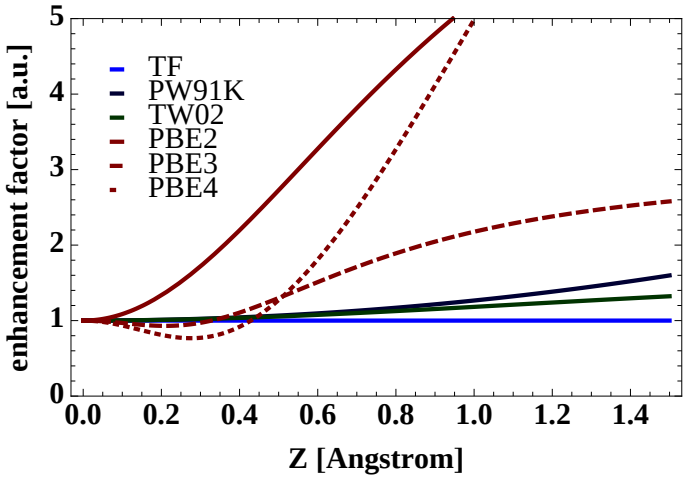


Figure 4.1: Enhancement functions used in kinetic energy functionals plotted in the domain $0.0 < s < 1.5$. To guide the eye we also display the constant value of 1 that corresponds to the original Thomas-Fermi ansatz.

used, and the supermolecular basis set expansion (denoted as FDE(s)),[72] in which the basis functions of both subsystems are used. If not stated otherwise, electron densities of both subsystems were relaxed and converged in six freeze-and-thaw cycles.[68] The approximate kinetic energy functionals for the non-additive kinetic energy used in this work are the Thomas-Fermi functional (TF),[108–110] the Thomas-Fermi plus 1/9 von Weizsäcker (TF9W) functional,[115] the PW91K functional,[72,116] the Tran-Wesolowski functional (TW02),[120] and the Karasiev-Trickey-Harris PBE2, PBE3, PBE4 functionals.[105] These PBEn kinetic-energy functionals were implemented in ADF for this work. For the purpose of analysis we also employed a purely electrostatic embedding in which the embedding potential contained only the Coulomb interaction with the frozen system, that is, omitting the non-additive kinetic energy and exchange-correlation energy contributions.

To setup and execute all our calculations and then to retrieve the data, we used PyADF[SMB6] which is a scripting framework for quantum chemistry implemented in the Python[129] programming language.

The electron density and the quantities derived from it were obtained on the integration grid used by ADF from a locally modified version of the DENSF-utility program of the ADF package. For visualization, the electron deformation density, the enhancement factor, and the effective embedding potential were obtained

on an evenly spaced grid by the same locally modified version of DENSF. The electron deformation density $\rho_{\mathrm{def}}(\boldsymbol{r})$ is defined as

$$\rho_{\mathrm{def}}(\boldsymbol{r}) = \rho_{\mathrm{SCF}}(\boldsymbol{r}) - \rho^{(1)}_{\mathrm{frag}}(\boldsymbol{r}) - \rho^{(2)}_{\mathrm{frag}}(\boldsymbol{r}) \tag{4.13}$$

where $\rho^{(i)}_{\mathrm{frag}}(\boldsymbol{r})$ (i = 1,2) is the electron density of an isolated fragment and $\rho_{\mathrm{SCF}}(\boldsymbol{r})$ the final, converged SCF electron density. In the case of FDE, we have $\rho_{\mathrm{SCF}}(\boldsymbol{r})= \rho^{(1)}_{\mathrm{SCF}}(\boldsymbol{r}) + \rho^{(2)}_{\mathrm{SCF}}(\boldsymbol{r})$ where $\rho_{\mathrm{SCF}}(\boldsymbol{r})= \rho^{\mathrm{FDE(6)}}_{\mathrm{total}} = \rho^{\mathrm{FDE(6)}}_{1} + \rho^{\mathrm{FDE(6)}}_{2}$. In order to quantitatively visualize the deformation density along the bond axis (z-axis), we adopted a scheme inspired by the one employed by Belpassi *et al.*[107] We numerically integrated on the evenly spaced grid used by DENSF (integration accuracy of 10^{-4} a.u. for the electron deformation density) over the x- and y-coordinates for every corresponding point on the z-axis,

$$\tilde{\rho}(z) \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho(x,y,z)dxdy, \tag{4.14}$$

where ρ(x, y, z) is the deformation density. This condenses the information contained in these three-dimensional functions to one-dimensional functions that are easier to plot. Since results of the calculations with the supermolecular basis set expansion were very similar, all the figures throughout this work display the data obtained with the default monomolecular basis set expansion. For each Ng–AuF (Ng=Ar, Kr, Xe) molecule, the calculations were done at a single geometry, with bond lengths taken from the analysis of experimental microwave data. The Ng–Au distances are d(Ar–Au) = 2.391 Å,[130] d(Kr–Au) = 2.461 Å[131] and d(Xe–Au) = 2.543 Å,[106] while the AuF distance was fixed at 1.918 Å.[107]

4.4 Results and Discussion

The Ng–Au bonds in NgAuF complexes are weakly covalent, with the interaction strength increasing from 49 to 94 kJ/mol upon moving from Ar to Xe.[107] In Section 4.1 we consider only the ArAuF molecule, but quantitative differences among NgAuF molecules are discussed in Section 4.2 and Section 4.4. Since the Ng–Au bond is much weaker than the Au–F bond (computed as 296 kJ/mol for the free molecule[107]), this naturally calls for a partitioning into Ng and AuF fragments in FDE with the Ng atom taken as the active subsystem and AuF taken as the frozen subsystem. When we consider the Ng–Au bond formation in this framework, electron density should move from the Ng atom into the bonding region (for details see Ref.[107]). This process can in principle be described by the FDE approach, possibly with the aid of the freeze-and-thaw update procedure to allow for some adjustment of the AuF density.

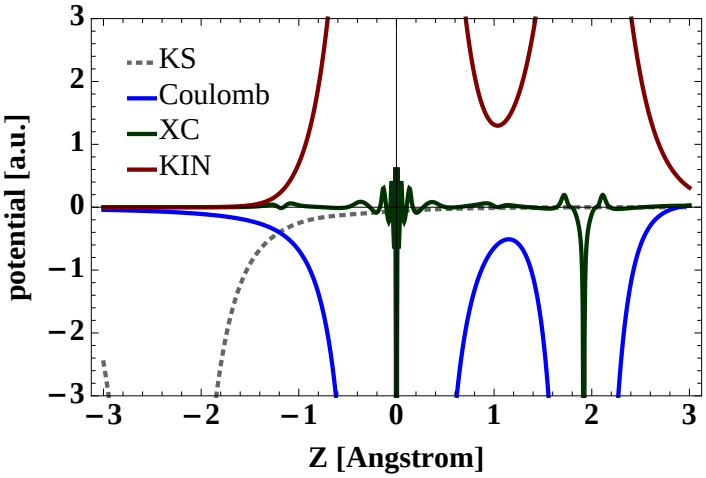


Figure 4.2: ArAuF: The Kohn-Sham potential and contributions of the embedding potential due to electrostatic (nuclear and Coulomb), non-additive exchange-correlation and kinetic energy for the TF kinetic energy functional along the bond axis. The Au atom is situated at $z = 0.0\,\text{Å}$, F at $z = 1.92\,\text{Å}$, and Ar at $z = -2.39\,\text{Å}$.

4.4.1 Embedding Potentials

For fixed densities and exchange-correlation functional, any difference in the embedding potential $v_{\text{eff}}^{\text{emb}}[\rho^{(1)}, \rho^{(2)}](\boldsymbol{r})$ is due only to differences in the approximate kinetic-energy functional that is used. We begin our analysis by taking the density of the isolated fragments to calculate the embedding potential $v_{\text{eff}}^{\text{emb}}$ as well as its components, [See Equation (4)]. This corresponds to the embedding potential as it is used in the first iteration of a self-consistent solution of the KSCED equations (1).

Figure 4.2 shows a plot of the components of $v_{\text{eff}}^{\text{emb}}(\boldsymbol{r})$ for the Ar atom along the bond axis for the TF kinetic-energy density functional. Figures 4.3 and 4.4 show the plots of the full embedding potential for the kinetic-energy functionals investigated in this work.

At large distances from the frozen AuF unit (for $z < -2.0$ Å) all embedding potentials are dominated by the attractive electrostatic contribution. The employed (semi-)local non-additive kinetic-energy and the non-additive exchange-correlation potentials approach zero because the derivatives in Equation 4.4 are evaluated at the same value of ρ when $\rho^{(2)}(\boldsymbol{r})$ is negligibly small. The other

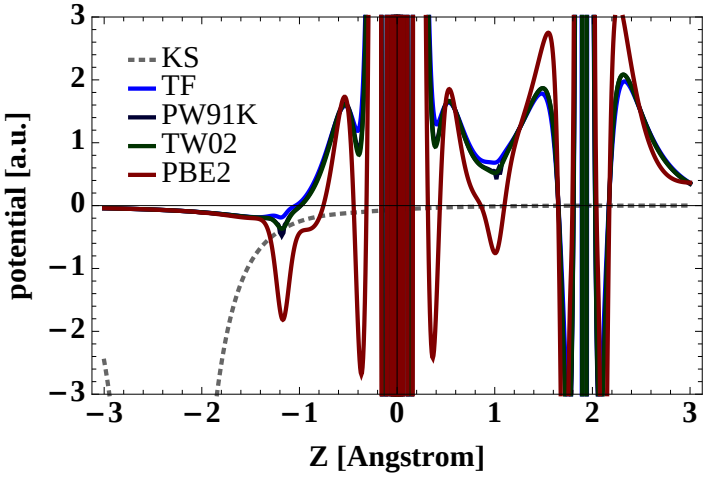


Figure 4.3: ArAuF: The Kohn-Sham and embedding potentials generated using the Thomas-Fermi (TF), PW91K, TW02 and PBE2 functionals along the bond axis. The Au atom is situated at $z = 0.0\,\text{Å}$, F at $z = 1.92\,\text{Å}$, and Ar at $z = -2.39\,\text{Å}$.

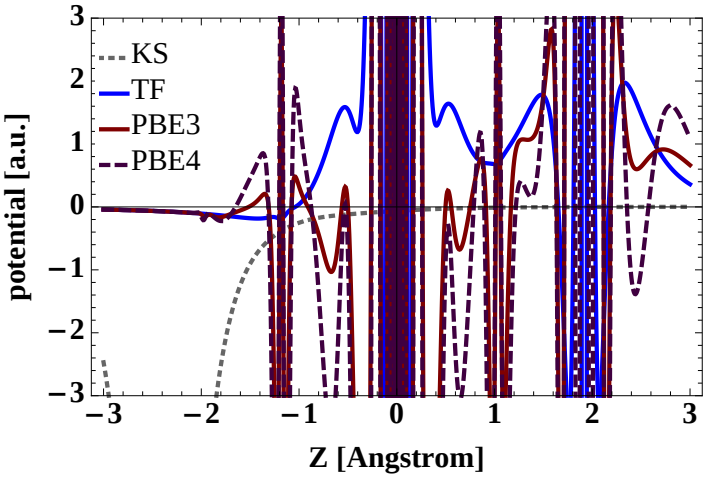


Figure 4.4: ArAuF: The Kohn-Sham potential and embedding potentials generated using the Thomas-Fermi (TF), PBE3 and PBE4 functionals along the bond axis. The Au atom is situated at $z = 0.0\,\text{Å}$, F at $z = 1.92\,\text{Å}$, and Ar at $z = -2.39\,\text{Å}$.

extreme is found in the vicinity of the frozen nuclei where we find an oscillatory potential resulting from competition between the large positive kinetic energy contribution and the strongly negative Coulomb potential. This highly oscillatory potential resulting from relatively smooth individual components (cf. Figure 4.2 for the components and the blue line in Figure 4.3 for the resulting TF embedding potential) is consistent with the assumption made in the FDE ansatz in which the active density should complement the frozen density to yield the exact total density. In the case of a heavy atom with a pronounced shell structure such oscillatory potentials should represent the Pauli repulsion that arises due to the frozen core orbitals. This core region is overall repulsive with the shallow negative regions outweighed by regions in which the total potential is strongly positive.

The region of interest is the area between $z = -2.0\,\text{Å}$ and $z = -0.5\,\text{Å}$ in which the charge transfer from the noble gas to the AuF is known to take place and where subtle differences between the kinetic-energy contributions yield an important contribution to the relatively shallow potential. To put this contribution in perspective we have also included the KS potential of the isolated Ar atom to which this embedding potential will be added in the plots. The TF curve climbs almost monotonically, with the repulsive wall starting at about -1.0 Å. The PW91K and TW02 functionals deepen the well somewhat and introduce some substructure. This is amplified by the PBEn functionals of which we shown only the PBE2 functional together with the other GGA functionals to not clutter Figure 4.3. This functional gives rise to a significant deepening of the well near the gold atom and also renders the potential in the region between the gold and the fluoride attractive. The use of higher-order terms in PBE3 and PBE4 is depicted in figure 4.4 in which we observe oscillations also in the bonding region of the complex.

To show the cause of these differences in embedding potentials more clearly Figure 4.5 shows the enhancement factor $F_\text{t}(s)$ for the total (sum of fragments) density of the complex, i.e. $\rho^\text{Ar}(\boldsymbol{r}) + \rho^\text{AuF}(\boldsymbol{r})$, for different approximations of the non-additive kinetic-energy functionals.

It is the derivative of the enhancement factor with respect to the density which determines the potential. The trends observed in the embedding potential are indeed visible in the enhancement factor plots, with the PW91K and TW02 curves varying smoothly compared to the PBE2 functional that differs only in parametrization but not in functional form from TW02. Interesting is the great similarity between the PW91K and TW02 curve that deviate only in the uninteresting high s regime that starts at the low density tail of the fluoride and argon (not visible in this plot). We do not show the enhancement factor of PBE3 and PBE4 because these exhibit strong oscillations and would clutter the figure. Already from the

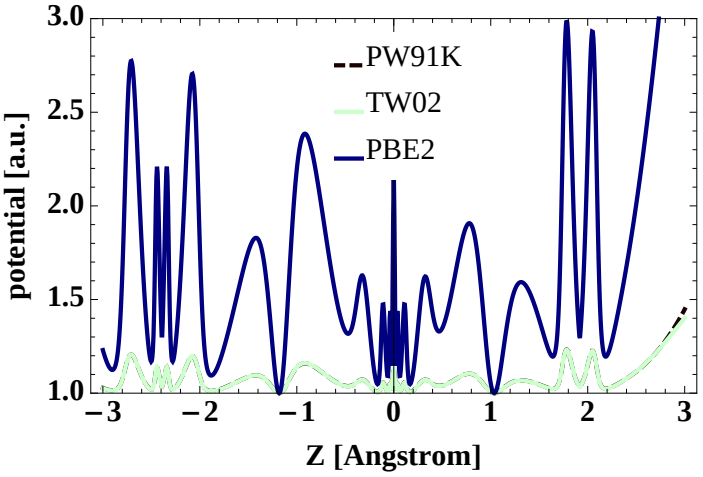


Figure 4.5: ArAuF: Enhancement factor for the total (sum of fragments) density i.e. $\rho^{\mathrm{Ar}} + \rho^{\mathrm{AuF}}$, for different approximations along the bond axis. The Au atom is situated at $z = 0.0$ Å, F at $z = 1.92$ Å, and Ar at $z = -2.39$ Å.

potential it is clear, however, that we may expect a stronger charge transfer for the PBEn functionals than for the other functionals.

4.4.2 Induced dipole moments

As a simple direct measure of the changes in the electron density we take the dipole moment changes that occur upon formation of the Ng–AuF (Ng=Ar, Kr, Xe) molecules. This property is ideally suited to assess the accuracy of the electron density that results from FDE calculations, both in comparison to the reference KS method as well as in relation to simple electrostatic models. For the latter we consider a simple model that neglects charge transfer and takes the electric field strength at the position of the Ng atom to compute an induced dipole moment, $\mu^{\mathrm{ind}} = \mathbf{E}\,\alpha$ using the relation between this electric field strength $\mathbf{E}$ and the dipole polarizability α of the atom. The experimental values for the polarizability of the noble gas atoms were used, as obtained from Ref.[132] This simple point polarizability model will be denoted (PP) and can also be compared with the electrostatic embedding model (EE) in which the non-additive kinetic energy and exchange-correlation contributions are completely neglected in the KSCED procedure. Table 1 compares the induced dipole moments of Ng (Ng=Ar, Kr, Xe) molecules calculated using these models, to the ones calculated with the reference KS and approximate FDE schemes.

Table 4.1: Magnitude of the induced dipole moments in Debye of Ar, Kr and Xe due to the interaction with the AuF molecule with different approximations. The KS induced dipole moment is obtained by subtracting the permanent dipole moment of AuF from the dipole moment of the Ng–AuF complex.

	Ar	Kr	Xe	Ar	Kr	Xe
KS				1.80	2.15	2.67
PP	0.10	0.14	0.20			

	FDE(m)			FDE(s)		
EE†)	1.43	5.81	8.41	—	—	—
TF	0.27	0.36	0.33	0.26	0.30	0.31
TF9W	0.37	0.54	0.61	0.39	0.49	0.58
PW91K	0.36	0.52	0.57	0.39	0.49	0.56
TW02	0.36	0.53	0.58	0.41	0.50	0.58
PBE2	1.09	1.99	2.61	1.78	2.26	2.88
PBE3†)	1.15	2.29	3.13	—	—	—
PBE4†)	1.72	3.99	5.46	—	—	—

†) FDE(s): SCF convergence could not be reached.

As expected, the induced dipole moment increases from Ar to Xe due to the larger polarizability that overcomes the greater bond distance. For a system which exhibits a considerable covalent character like Xe–AuF, the PP model clearly predicts a too small effect. Missing in the PP model is the charge transfer from the Ng atom to the AuF unit, as explained in detail by Belpassi *et al.*[107] Since FDE allows for such charge transfer, we may check how much FDE does improve upon the classical model. This cannot lead to perfect agreement as there is a small region close to the gold nucleus in which the AuF frozen density exceeds that of the NgAuF complex (the integrated negative density is -0.051, -0.064 and -0.084, for Ng=Ar, Kr and Xe, respectively). Adjustment of the AuF density by freeze-thaw cycles to allow for this effect will be considered in the next section, but for the present qualitative purpose it is reasonable to assume that these small regions of negative density do not influence the result too much. This is also consistent with the experimental evidence that the Ng–Au bonding hardly influences the strong AuF bond much.[106,130,131] As can be seen from Table 1, the TF, TF9W, PW91K and TW02 functionals underestimate the dipole moment change for all

noble gases in the series, Ar, Kr, Xe but do improve upon the PP model. The underestimation is not due to basis set deficiencies as calculation with the FDE(s) approach (in which the supermolecular basis is used) yields values close to the FDE(m) values. This indicates that the embedding potential generated with these functionals is too shallow, overestimating the Pauli repulsion that arises from the AuF unit.

The deeper well generated by PBE2 and PBE3 leads to dipole moment changes that approach the results of supermolecular KS calculations. The PBE2 functional give values close to the reference result if the FDE(s) expansion, in which the basis set is equal to the one used in the reference KS calculation, is used. On the other hand one may observe that the more complicated embedding potential generated with the PBE4 functional is too attractive, leading to dipole moment changes that are well above those obtained by the KS method. Removing the repulsive non-additive KE and XC terms by considering only EE results in a large overestimation of the charge transfer. These values provide an indication of the upper limit of charge transfer that is possible within the monomolecular expansion used in the FDE(m) approach. For the EE model, as well as for the PBE3 and PBE4 functionals it was not possible to reach convergence in the FDE(s) expansion that allows full charge transfer to the AuF unit.

4.4.3 Deformation Densities

We now allow for adjustment of the AuF density as well and define deformation densities as explained in Section 4.3. After six freeze-and-thaw cycles, the electron densities hardly change, and the quantities discussed here are converged within the reported accuracy. In order to assess the performance of these FDE deformation densities we compared them with the KS deformation density defined as $\rho^{\mathrm{KS}}(\boldsymbol{r})$ - $\rho^{(1)}_{\mathrm{frag}}(\boldsymbol{r}) - \rho^{(2)}_{\mathrm{frag}}(\boldsymbol{r})$, where $\rho^{\mathrm{KS}}(\boldsymbol{r})$ designates the density of the supermolecule obtained with the conventional KS method. These deformation densities for Ar–AuF, radially integrated and plotted along the bond axis (See Section 2 for details) are presented in Figure 6. The deformation density obtained for the PBE4 functional is omitted since it significantly deviates from the KS and that it has a complicated structure.

In the KS deformation density, there is a sharp increase at the very center of the noble gas delimited on both sides by a decrease. The increase in density at the noble gas nucleus is due to the empty 4s orbital of Ar that participates in the molecular orbitals of the complex, resulting in a larger s-orbital occupation on the

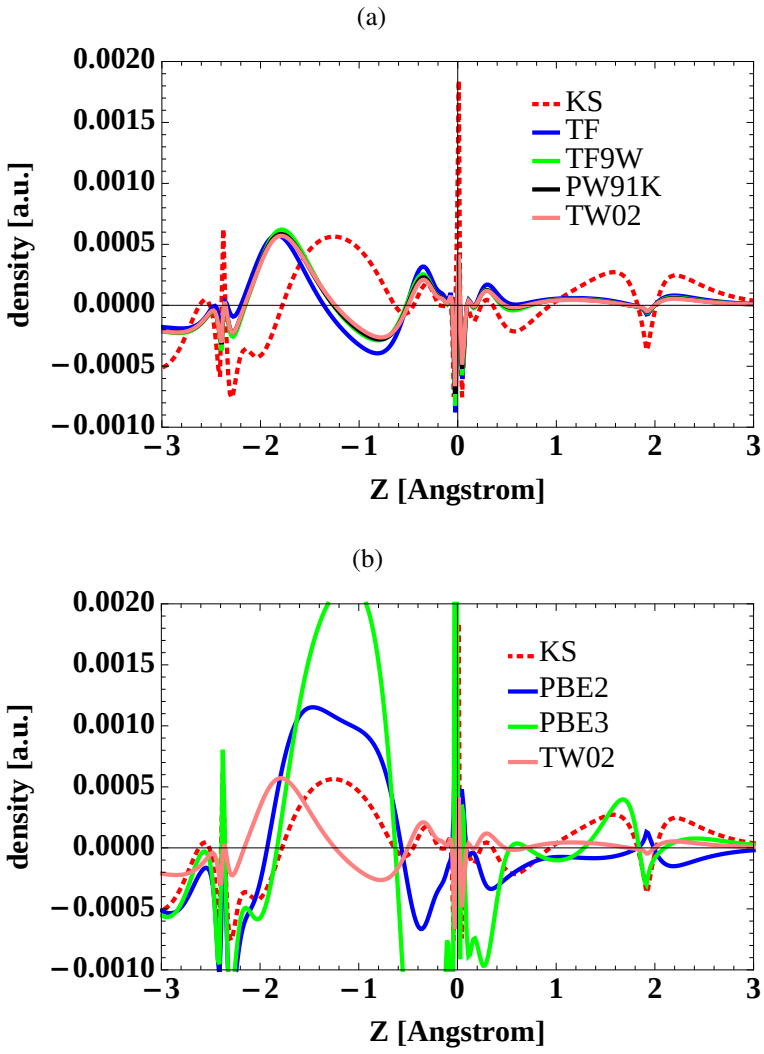


Figure 4.6: Radially integrated deformation densities for different approximations of the non-additive kinetic-energy functional for the ArAuF molecule. KS denotes the difference between the KS supermolecular density and the one of sum of fragments. FDE deformation densities are calculated by taking the difference between the KSCED and sum of fragments density. The Au atom is situated at $z = 0.0$ Å, F at $z = 1.92$ Å, and Ar at $z = -2.39$ Å.

noble gas. The decrease on both sides is caused by donation from the occupied 3p-orbital that participates in the weak sigma bond to the gold center. Together, this corresponds to an overall charge transfer from Ar to AuF and decrease of the density near the Ar nucleus. The TF, TF9W, PW91K and TW02 functionals all capture this trend but underestimate its magnitude with a maximum of charge build-up in the bonding region at $z = -1.8$ Å, rather than at at $z = -1.1$ Å as in the KS-reference. The PBE2, PBE3 and PBE4 functionals overestimate the magnitude of charge transfer. Among all functionals, PBE2 comes closest to the KS results if we look only at the density changes near Ar and in the bonding region. The picture changes when considering also the density changes near the Au center. TF, TF9W, PW91K and TW02 now resemble the KS picture, although again underestimating the magnitude of the density distortion. The PBE2, PBE3 and PBE4 (not shown on the picture) functionals show a too complicated density deformation. Finally, in the vicinity of the fluorine nucleus all the functionals tested here, except for PBE2, capture qualitatively the trends. Again the first family of functionals (TF, TF9W, PW91K, TW02) underestimates the density changes, while the PBE3 functional has a very good agreement with the KS result and that the PBE4 functional (not shown on the picture) overestimates it.

4.4.4 Quantitative Analysis

To quantify the performance of the approximate non-additive kinetic-energy functionals under consideration, we used three measures of accuracy (see Ref. 87,133), again using the supermolecular KS method as a reference and considering the density after six freeze-and-thaw cycles. The measures are the dipole moment difference for the total system,

$$\Delta\mu = \int \left(\rho^{\mathrm{KS}}(\boldsymbol{r}) - \left(\rho_1^{\mathrm{FDE}(6)}(\boldsymbol{r}) + \rho_2^{\mathrm{FDE}(6)}(\boldsymbol{r})\right)\right)\boldsymbol{r}d\boldsymbol{r}; \tag{4.15}$$

the integrated absolute difference density,

$$\Delta^{\mathrm{abs}} = \frac{1}{N}\int \left|\left(\rho^{\mathrm{KS}}(\boldsymbol{r}) - \left(\rho_1^{\mathrm{FDE}(6)}(\boldsymbol{r}) + \rho_2^{\mathrm{FDE}(6)}(\boldsymbol{r})\right)\right)\right| d\boldsymbol{r}; \tag{4.16}$$

and the root-mean-square error in the density,

$$\Delta^{\mathrm{rms}} = \frac{1}{N}\sqrt{\int \left(\rho^{\mathrm{KS}}(\boldsymbol{r}) - \left(\rho_1^{\mathrm{FDE}(6)}(\boldsymbol{r}) + \rho_2^{\mathrm{FDE}(6)}(\boldsymbol{r})\right)\right)^2 d\boldsymbol{r}}. \tag{4.17}$$

Here, $\rho^{\mathrm{KS}}(\boldsymbol{r})$ is the density of the supermolecule obtained with the conventional KS method, $\rho_1^{\mathrm{FDE}(6)}(\boldsymbol{r})$ and $\rho_2^{\mathrm{FDE}(6)}(\boldsymbol{r})$ are the densities obtained with the FDE scheme, and N is the total number of electrons of the systems under investigation. These data are presented in Table 2 for the molecules Ar–AuF, Kr–AuF and Xe–AuF, respectively. The "sum of fragments" designates the superposition of the densities obtained by KS calculations on the isolated fragments (Ng and AuF). This is the starting density on which the FDE scheme is expected to improve upon. The dipole moments of AuF, Ar–AuF, Kr–AuF and Xe–AuF are 3.61 D, 5.47 D, 5.76 D and 6.28 D, respectively. It can be seen that the relative accuracy drops when one moves from Ar to Xe for all functionals, except for PBE2 and PBE3 when going from Ar to Kr. This overall loss in accuracy is expected given the increase of covalent character of the bond between the Ng and AuF upon moving from Ar to Xe. This drop in accuracy is hardly visible for Δ^{rms} because, compared to Δ^{abs}, this measure puts more weight on the nuclear regions of gold where the density and error does not change so much through the series.

As expected, TF, TF9W, PW91K and TW02 functionals all improve the sum of fragments starting density, although the magnitude of this improvement is small. Considering only the dipole moment difference, the TF9W model appears to yield the best results. Regarding the other two measures of accuracy, the TW02 functional performs best. In general the differences in results between these functionals are rather small, especially compared to the absolute size of the errors that amounts to more than 1 Debye in the dipole moment. The popular PW91K functional gives in this case less accurate results than TF9W and TW02.

It is noteworthy that for PBEn functionals, results with six freeze-thaw cycles shown in Table 2 are worse than one would expect from the results without freeze-thaw cycles shown in Table 1. This is due to significant changes in the AuF density during freeze-and-thaw cycles. Among others, the dipole moment of the AuF unit changes significantly (1.02, 0.91 and 1.00 D, for ArAuF, KrAuF, XeAuF, respectively) when employing these functionals contrary to what is to be expected on basis of the weak intermolecular interaction and the strong intramolecular interaction. After freeze-and-thaw cycles the PBE2 and PBE3 functionals now yield worse results than the other conjointness functionals, with PBE2 only slightly improving over the sum of fragments. The larger departure from the TF reference used in the PBEn functionals does thus lead to artifacts when the frozen density is allowed to relax. This is also the cause of the convergence problems observed in the FDE(s) calculations in which the full AuF basis is available in the noble gas subsystem calculation. A similar behavior has been observed in our earlier work for other functionals,[SMB4, SMB2] among which also the PW91K functional. This can be remedied to some extent by either cutting off the core regions of the poten-

tial using a switching function[SMB4] or by allowing for a kinetic energy functional that is no longer decomposable into two separate contributions from the total and active density.[84] As both approaches are only valid in the limit of non-overlapping densities we will not consider them here in detail. Exploratory calculations with the switching function used in reference[SMB4] do not show improvement of results obtained with PBE2 after freeze-and-thaw, increasing the underestimation of the dipole moment from 1.74 to 1.87 Debye. This was expected since the switching function is constructed to test for a small overlap situation and will not fully activate the correction in the current case.

The PBE3 and PBE4 functionals are not suitable for use in a full optimization scheme that includes freeze-and-thaw relaxation cycles. For these functionals all measures of accuracy deteriorate as compared to the "sum of fragments" starting density. Also for PBE2, while still yielding an acceptable final result, the use of freeze-thaw cycles will lead to detoriation rather than an improvement of the promising results (cf. Table 1) that are obtained without this relaxation.

4.5 Conclusions

The effective-embedding potentials produced by the currently available kinetic-energy density functionals for use in the FDE method appear to be too repulsive to properly describe the coordination bonding found in the noble gas atom bound to the gold fluoride molecule. The use of increasing orders of the reduced density gradient s in the expression of GGA kinetic-energy functionals does not necessarily yield improved functionals as can be seen from the bad performance of the PBE4 functional. This could, however, be an effect of the chosen parametrization that was developed for a quite different purpose and bonding situation. The performance of a given GGA functional varies significantly depending on the location on the bond axis. For example, the density obtained with the PBE3 functional follows the KS trend closely near the Ar and F atoms whereas in the bonding region and near the Au atom it fails.

The general conclusion is that for systems like NgAuF (Ng=Ar, Kr, Xe), none of the enhancement functions used in the approximate kinetic-energy functionals is yet able to describe the weak covalent bond adequately. The PBE2 functional comes closest to the Kohn-Sham result, but can only be used with a frozen AuF unit as updating the AuF density worsens the result. But also functionals with a proven accuracy in other bonding situations, like PW91K, fail for this more challenging bonding situation. It is important to note that GGA functionals such as

PW91k are able to accurately describe much stronger bonds, such as the hydrogen bond in F–H–F$^-$.[82] Therefore, the bond strength alone is not a sufficient criterion to judge whether GGA functionals can be expected to yield an adequate description, one also needs some a priori knowledge of the bonding characteristics.

The failure of GGA kinetic-energy functionals even for weak covalent bonds raises the question as to whether the "conjointness" approach as applied in previous work is suitable to derive functionals that can describe the stronger interactions which are of interest in many applications of subsystem methods.[134] It could be interesting to approximate the kinetic-energy component of the embedding potential directly in a non-decomposable fashion[SMB484] also in regions in which the frozen and active densities show significant overlap. An alternative is to introduce approximations that not only locally depend on the electron densities of the subsystems, but that also depend on the KS orbitals of the subsystems. This would make the method similar to a pseudopotential approach in which the orbital information is also used to model the effect of the frozen density.

Table 4.2: The dipole moment difference $\Delta\mu$ in Debye, the integrated absolute difference density Δ^{abs} and root-mean-square error in the density Δ^{rms} for Ar–AuF, Kr–AuF and Xe–AuF. All the data refer to the comparison to KS-DFT. All FDE calculations are done in the monomolecular expansion FDE(m).

	Ar–AuF			Kr–AuF			XeAuF		
	$\Delta\mu$	$\Delta^{\mathrm{abs}}\times10^3$	$\Delta^{\mathrm{rms}}\times10^3$	$\Delta\mu$	$\Delta^{\mathrm{abs}}\times10^3$	$\Delta^{\mathrm{rms}}\times10^3$	$\Delta\mu$	$\Delta^{\mathrm{abs}}\times10^3$	$\Delta^{\mathrm{rms}}\times10^3$
Sum of Fragments	1.80	4.04	0.41	2.15	4.29	0.39	2.67	4.61	0.41
TF	1.04	3.87	0.44	1.24	4.07	0.43	1.86	4.67	0.45
TF9W	1.07	3.48	0.40	1.11	3.59	0.39	1.59	4.23	0.42
PW91K	1.10	3.52	0.39	1.19	3.63	0.39	1.75	4.30	0.41
TW02	1.12	3.49	0.39	1.22	3.59	0.38	1.79	4.22	0.41
PBE2	1.74	4.94	0.47	0.97	4.25	0.41	1.25	4.83	0.44
PBE3	2.36	9.48	1.00	2.45	9.43	0.97	7.16	14.98	1.32
PBE4	7.46	26.12	2.83	8.13	24.49	2.09	12.33	25.91	2.25

Chapter 5

Constructing an Accurate Embedding Potential for a Given Density

In this Chapter, an accurate potential reconstruction scheme for a given electron density is presented. This scheme is used to investigate the validity of the freeze-and-thaw method. Two model systems with covalent bonding: LiH and Li_2 are used, partitioned into Li^+ and H^-; Li^+ and Li^- subsystems; respectively. The advantage of such a partitioning is that it provides two strongly different starting points for these simple model systems, for which one can analyse the effect of freeze-and-thaw cycles. The effect of freeze-and-thaw cycles on the quality of reconstructed embedding potentials and electron densities is analysed. Furthermore, for LiH, given that the system is v_s-representable, the exact potential for the 2-electron H^- fragment is calculated, which is compared with that obtained from the reconstruction scheme.

5.1 Potential Reconstruction

For the subsystems that interact covalently, the currently available approximations of $v_{\mathrm{T}}^{\mathrm{nadd}}[\rho^{(1)}, \rho^{(2)}]$ fail.[134] For example, recent studies have shown that description of the coordination bond between ammonia borane[83] as well as the weak covalent bond between noble gas atoms and AuF[SMB5] fails. Available approximations also fail for transition metal complexes such as carbonyl compounds where π-bonding

plays an important role.[83]

It is of interest to treat subsystems connected by covalent bonds, for example for a subsystem description of biological systems. Recently, a scheme where capping groups are introduced to model covalent bonds between amino acid building blocks of proteins has been proposed.[87] However, this scheme does not solve the insufficiency and problems of the currently available approximations of $v_{\mathrm{T}}^{\mathrm{nadd}}[\rho^{(1)}, \rho^{(2)}]$ but instead, it provides a pragmatic solution. Therefore, improved approximations of $v_{\mathrm{T}}^{\mathrm{nadd}}[\rho^{(1)}, \rho^{(2)}]$ must be developed to treat biological systems. To achieve this, one can use the same strategy that is successfully employed in the development of approximations to the exchange-correlation potential, which led, for instance to the SAOP potential[44,135,136] and the KTn family of exchange-correlation functionals.[137,138] In this strategy, one investigates the exact properties of the exchange-correlation potential and deficiencies of the available approximations by using the knowledge of accurate reference potentials for selected systems. This knowledge, therefore, guides to construct improved approximations. Exact reference potentials for $v_{\mathrm{T}}^{\mathrm{nadd}}[\rho^{(1)}, \rho^{(2)}]$ are already known for special limits such as infinitely separated systems[SMB4] or close to the nuclei of the frozen subsystems[84] and for four-electron systems.[139]

Potential Reconstruction Recently, Fux *et al.*[86] have implemented a method for the numerical calculation of accurate references to $v_{\mathrm{T}}^{\mathrm{nadd}}[\rho^{(1)}, \rho^{(2)}]$. They obtained accurate embedding potentials for a set of model systems where subsystems are connected by hydrogen bonds of various strength, a coordination bond and a covalent bond, which they compared to those obtained from popular kinetic-energy density functionals.

In their implementation, they chose to use the direct optimisation method of Wu and Yang[140] with an additional constraint which ensures that the obtained potentials are smooth, as proposed by Bulat *et al.*[141] The method of Wu and Yang first defines a Lagrangian functional $W_s[v_s(\boldsymbol{r})]$ depending on the sought potential $v_s(\boldsymbol{r})]$

$$W[v_s(\boldsymbol{r})] = 2\sum_{i}^{\frac{N^{(1)}}{2}} \langle\phi_i|\hat{T}|\phi_i\rangle + \int v_s(\boldsymbol{r})[\rho_{\mathrm{rec}}^{(1)}(\boldsymbol{r}) + \rho^{(2)}(\boldsymbol{r}) - \rho^{\mathrm{ref}}(\boldsymbol{r})]\mathrm{d}^3\boldsymbol{r}, \quad (5.1)$$

where ρ^{ref} is the reference density, i.e. the total electron density obtained from a KS-DFT calculation on the full system, $\rho_{\mathrm{rec}}^{(1)}(\boldsymbol{r})$ and $\{\phi_i(\boldsymbol{r})\}$ denote the reconstructed electron density and the Kohn-Sham orbitals that result from $v_s(\boldsymbol{r})$; respectively, and $\rho^{(2)}(\boldsymbol{r})$ is the electron density of the frozen subsystem 2. If the

Lagrangian functional is maximised, $W[v_s(\boldsymbol{r})]$ equals the kinetic-energy $T_s[\rho^{\mathrm{ref}}]$ and $v_s(\boldsymbol{r})$ is the sought potential. The potential $v_s(\boldsymbol{r})$ is expanded into an initial guess $v_{\mathrm{initial}}(\boldsymbol{r})$ and a linear combination of auxiliary functions $g_t(\boldsymbol{r})$ that models the difference between the initial guess and the final potential,

$$v_s(\boldsymbol{r}) = v_{\mathrm{initial}}(\boldsymbol{r}) + \sum_t b_t g_t(\boldsymbol{r}). \tag{5.2}$$

The first and second derivatives of $W[v_s(\boldsymbol{r})]$ with respect to the expansion coefficients $\{b_t\}$ can be calculated analyticaly and $W[v_s(\boldsymbol{r})]$ is maximised without any constraints.

It is a well known problem that the reconstruction of a potential from a given electron density does not always result in a final potential that is physically meaningful. As a result of the use of a finite basis set expansion that represents the KohnSham orbitals and the electron density, infinitely many potentials may exist that reproduce the input electron density. This basis set expansion allows variations in the potential which causes changes in the reconstructed potential that cannot be detected by the electron density. This results in a reconstructed potential that has strong oscillations (for examples see Refs. 142–146). To overcome this problem, Fux *et al.*[86] chose to use the same strategy as in Ref. 142 and introduced a regularisation norm $||\nabla v_b(\boldsymbol{r})||$ in the Lagrangian functional to ensure smoothness of the reconstructed potentials,

$$\hat{W}_s[v_s(\boldsymbol{r})] = W_s[v_s(\boldsymbol{r})] + \lambda||\nabla v_b(\boldsymbol{r})||^2, \tag{5.3}$$

with

$$||\nabla v_b(\boldsymbol{r})||^2 = \sum_t \sum_u b_t b_u \langle g_t(\boldsymbol{r})|\nabla^2|g_u(\boldsymbol{r})\rangle, \tag{5.4}$$

where the squared norm of the gradient of the expansion-dependent term of the potential is minimised. It is of importance to obtain the smallest possible value for λ, which results in a smooth potential, such that among equivalent potentials, the smoothest one that gives the correct target electron density is chosen.

They developed a procedure for choosing λ where they initially perform a calculation using $\lambda = 0$ to determine the minimal error in the reconstructed electron density, measured as

$$\Delta_{\mathrm{abs}} = \int |\rho^{\mathrm{ref}}(\boldsymbol{r}) - \rho^{(1)}_{\mathrm{rec}}(\boldsymbol{r}) - \rho^{(2)}(\boldsymbol{r})|^2 \mathrm{d}^3\boldsymbol{r}. \tag{5.5}$$

Then, they start with a large value of λ, e.g. 10^{-3}, and successively decrease λ until the error in the reconstructed electron density $\rho^{(1)}_{\mathrm{rec}}(\boldsymbol{r})$ is smaller than 1.2 times the error when applying $\lambda = 0$.

Calculation of Accurate $v_{\mathrm{T}}^{\mathrm{nadd}}[\rho^{(1)},\rho^{(2)}]$ Once the sought potential $v_s(\boldsymbol{r})$, which equals the effective potential $v_{\mathrm{eff}}^{\mathrm{KSCED}}(\boldsymbol{r})$, is reconstructed, one can easily extract the nonadditive kinetic-energy component $v_{\mathrm{T}}^{\mathrm{nadd}}[\rho^{(1)},\rho^{(2)}]$. Using the relations in Eqs. (6.8) and (4.4), we can rewrite $v_{\mathrm{T}}^{\mathrm{nadd}}[\rho^{(1)},\rho^{(2)}]$ as

$$\begin{aligned} v_{\mathrm{T}}^{\mathrm{nadd}}[\rho_{\mathrm{rec}}^{(1)},\rho^{(2)}] &= v_s(\boldsymbol{r}) - v_{\mathrm{eff}}^{\mathrm{KS}}[\rho_{\mathrm{rec}}^{(1)}](\boldsymbol{r}) - v_{\mathrm{nuc}}^{(2)}(\boldsymbol{r}) \\ &\quad - \int \frac{\rho^{(2)}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\mathrm{d}^3\boldsymbol{r}' - \mathrm{v}_{\mathrm{xc}}^{\mathrm{nadd}}[\rho_{\mathrm{rec}}^{(1)},\rho^{(2)}], \end{aligned} \tag{5.6}$$

where $\rho_{\mathrm{rec}}^{(1)}(\boldsymbol{r})$ is the reconstructed electron density of the active subsystem 1, i.e. $\rho^{(1)}(\boldsymbol{r})$. We can further write this as

$$v_{\mathrm{T}}^{\mathrm{nadd}}[\rho_{\mathrm{rec}}^{(1)},\rho^{(2)}] = v_s(\boldsymbol{r}) - v_{\mathrm{ext}}(\boldsymbol{r}) + \int \frac{\rho^{(\mathrm{tot})}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\mathrm{d}^3\boldsymbol{r}' - \mathrm{v}_{\mathrm{xc}}[\rho^{(\mathrm{tot})}], \tag{5.7}$$

that is by subtracting the sum of the nuclear potential, electronic Coulomb potential, and the approximate exchange-correlation of the full system from the reconstructed potential.

In this work, we use this above presented scheme and built up on the work of Fux *et al.*[86] by investigating the validity of the freeze-and-thaw scheme in conjunction with the potential reconstruction. We analyze the effect of freeze-and-thaw cycles, the necessity of which was investigated earlier,[69] on the quality of reconstructed embedding potentials and electron densities. We use two model systems with covalent bonding: LiH and Li_2. For LiH, given that the system is v_s-representable, we calculate the exact potential for the 2-electron H^- fragment, which we compare with that obtained from the reconstruction scheme.

This work is organised as follows. We present the computational details in Section 5.2, the discussion of our results in Section 5.3 and finally the concluding remarks in Section 5.4.

5.2 Computational Details

All calculations were carried out using the FDE implementation[63,124] in the Amsterdam density functional package ADF.[126] All molecular structures were optimized with the PBE exchange-correlation functional[122] in combination with the QZ4P basis set of the ADF basis set library.[126]

In all FDE calculations, the supermolecular basis set expansion,[72] in which the basis functions of both subsystems are employed to expand the subsystem

electron densities, was used, and six freeze-and-thaw cycles were employed. We denote the use of supermolecular basis set expansion together with six freeze-and-thaw cycles as FDE(s)(6). TF9W kinetic-energy functional[115] has been applied to approximate the kinetic-energy component of the effective embedding potential.

For the potential reconstruction, we used the implementation of Fux *et al.*[86] in a locally modified version of ADF together with PYADF[SMB6] a scripting framework for quantum chemistry. As initial guess for $v_{\mathrm{initial}}(\boldsymbol{r})$, we used the potential obtained from a calculation of the isolated subsystem. We denote the use of freeze-and-thaw cycles with the potential reconstruction as $\mathrm{rec(i)}$ where i denotes the corresponding number of freeze-and-thaw cycles. We started potential reconstruction calculations with isolated fragments densities as well as with FDE(s)(6) densities. The former allows a direct comparison with the approximate FDE since the input densities are the same in both cases (i.e. isolated fragments densities). However, the latter gives an insight on the effect of starting from more accurate electron densities (i.e. FDE(s)(6) densities).

For the visualisation of the potentials, MATHEMATICA[147] was employed.

5.2.1 Choice of the Frozen Density

If the exact $v_{\mathrm{T}}^{\mathrm{nadd}}[\rho^{(1)}, \rho^{(2)}]$ is known, FDE yields the correct $\rho^{(1)}(\boldsymbol{r})$ density, such that $\rho^{(\mathrm{tot})}(\boldsymbol{r}) = \rho^{(1)}(\boldsymbol{r}) + \rho^{(2)}(\boldsymbol{r})$ is equal to the electron density that is obtained in a KS-DFT calculation on the full system. In order to achieve this, it is mandatory that the frozen $\rho^{(2)}(\boldsymbol{r})$ is chosen such that the exact $\rho^{(1)}(\boldsymbol{r}) = \rho^{(\mathrm{tot}}(\boldsymbol{r}) - \rho^{(2)}(\boldsymbol{r})$ fulfils two conditions.[60] First, $\rho^{(1)}(\boldsymbol{r})$ must be non-negative everywhere in space, i.e., $\rho^{(2)}(\boldsymbol{r}) \leq \rho^{(\mathrm{tot}}(\boldsymbol{r})$. This is because it is not possible to achieve the required $\rho^{(1)}(\boldsymbol{r})$ if it is negative in certain regions. Second, $\rho^{(1)}(\boldsymbol{r})$ must be a v_s-representable density, i.e. it must be possible to obtain this electron density as the ground state of a KohnSham potential. This condition cannot be fulfilled if the required $\rho^{(1)}(\boldsymbol{r})$ has nodal surfaces (i.e. where the density is zero). At such nodal surfaces, the potential would have to be singular [$v_{\mathrm{T}}^{\mathrm{nadd}}(\boldsymbol{r}) = \infty$], which is very difficult to represent in practical calculations. Common choices of $\rho^{(2)}(\boldsymbol{r})$, usually, do not fulfil these two conditions.

Finding a suitable partitioning of the total electron density into subsystem electron densities is, therefore, of great importance to calculate accurate reference potentials for $v_{\mathrm{T}}^{\mathrm{nadd}}(\boldsymbol{r})$. In this work, we try to overcome this issue by partitioning our two model systems $\mathrm{Li_2}$ and LiH into $\mathrm{Li^+}$ and $\mathrm{Li^-}$; $\mathrm{Li^+}$ and $\mathrm{H^-}$ subsystems; respectively. The advantage of such a partitioning is that it reduces the negativity of

$\rho^{(1)}(\boldsymbol{r})$ and, hence, is a great attempt to fulfil the first condition described above. Furthermore, it provides two strongly different starting points for these simple model systems for which one can analyse the effect of freeze-and-thaw cycles. For example, in the case of Li_2, the negative density, i.e. integrated negative part of the difference $\rho^{(\mathrm{tot}}(\boldsymbol{r}) - \rho^{(\mathrm{Li}^+)}(\boldsymbol{r})$, is only 0.001 whereas that of the difference $\rho^{(\mathrm{tot}}(\boldsymbol{r}) - \rho^{(\mathrm{Li}^-)}(\boldsymbol{r})$ is 0.852, where $\rho^{(\mathrm{Li}^+)}(\boldsymbol{r})$ and $\rho^{(\mathrm{Li}^-)}(\boldsymbol{r})$ are the electron densities obtained from a KS-DFT calculation on the isolated Li^+ and Li^- fragments; respectively. This naturally calls for choosing Li^+ as the frozen subsystem. Similarly, in the case of LiH, the negative density is 0.004 for $\rho^{(\mathrm{tot})}(\boldsymbol{r}) - \rho^{(\mathrm{Li}^+)}(\boldsymbol{r})$, and 0.295 for $\rho^{(\mathrm{tot})}(\boldsymbol{r}) - \rho^{(\mathrm{H}^-)}(\boldsymbol{r})$, and therefore, Li^+ is chosen as the frozen subsystem.

5.3 Results and Discussion

5.3.1 LiH

The LiH bond length is taken as $1.59\,\text{Å}$, with the H^- atom situated at $z = 0.80\,\text{Å}$, Li^+ at $z = -0.80\,\text{Å}$, where z is the bond axis. In all the plots H^- is the active subsystem.

Assessment of the Density To quantify the effect of freeze-and-thaw cycles, we used two measures of accuracy using the electron density that is obtained in a KS-DFT calculation on the full system as a reference. These measures are the dipole moment difference for the total system $\Delta\mu$ (Eq. (4.15)) similar to our earlier work,[SMB5] and the integrated difference density, Δ_{abs} (Eq. (5.5)) that is also used for choosing λ within the potential reconstruction scheme. These data are presented in Table 5.1, whereas, the effect of freeze-and-thaw cycles on the negative density (see Sec. 5.2.1 for its definition) is presented in Table 5.2. The s.o.f. designates the superposition of the densities obtained by KS calculations on the isolated fragments (H^- and Li^+). This is the starting density on which the FDE scheme is expected to improve upon. In contrast, TF9W fails to improve upon s.o.f. (Table 5.1). However, it results in a smaller negative density for both subsystems (Table 5.2).

Potential reconstruction scheme started with the s.o.f. densities improves the quality of the electron density upon TF9W as well as s.o.f.. This improvement is consistently present in successive freeze-and-thaw cycles. However, from rec(8)

to rec(10), the difference between successive cycles gets smaller (1.0E-4). This trend is also present in the negative densities (see Table 5.2).

On the other hand, potential reconstruction scheme starting from TF9W densities results in a higher quality of the electron density, with respect to that of starting from s.o.f., already in rec(0). rec(1) results in a further improvement, particularly, the dipole moment difference is 10.4 times smaller than that of rec(1) started with s.o.f. densities. This is also the cycle that results in the smallest negative density of ρ^{H^-} (1.26E-09). From rec(2) to rec(10), the negative density of ρ^{H^-} continues to increase, whereas the negative density of ρ^{Li^+} continues to decrease until rec(5), and from rec(5) to rec(10) it increases as well. It is noteworthy that even though the TF9W densities have a lower accuracy than those of s.o.f., if the potential reconstruction scheme starts from TF9W densities, it results in a higher quality of the reconstructed electron densities, and at an earlier cycle. This can be attributed to smaller negative density of both subsystems, which provides a more suitable starting point, as discussed in Section 5.2.1.

Table 5.1: LiH: Effect of freeze-and-thaw cycles on the the dipole moment difference for the total system $\Delta\mu$ (Debye), and the integrated difference density Δ_{abs} (Eq. (5.5)). s.o.f. denotes the sum of isolated fragments. Potential reconstruction calculations were started from the sum of isolated fragments as well as converged TF9W densities.

	Δ_{abs}	$\Delta\mu$	Δ_{abs}	$\Delta\mu$
s.o.f.	0.6321	2.0439	0.6321	2.0439
TF9W	0.6409	2.9345	0.6409	2.9345
	starting: s.o.f.		starting: TF9W	
rec(0)	0.0654	0.0520	0.0480	0.0629
rec(1)	0.0560	0.0698	0.0394	0.0067
rec(2)	0.0544	0.0736	0.0385	0.0154
rec(3)	0.0536	0.0743	0.0351	0.0288
rec(4)	0.0530	0.0749	0.0279	0.0153
rec(5)	0.0527	0.0748	0.0337	0.0320
rec(6)	0.0524	0.0746	0.0397	0.0721
rec(7)	0.0522	0.0746	0.0426	0.0756
rec(8)	0.0520	0.0746	0.0446	0.0767
rec(9)	0.0519	0.0748	0.0461	0.0750
rec(10)	0.0518	0.0747	0.0471	0.0732

Table 5.2: LiH: Effect of freeze-and-thaw cycles on the negative density, i.e. integrated negative part of the difference $\rho^{(\mathrm{tot})}(\boldsymbol{r}) - \rho_i^{(2)}(\boldsymbol{r})$ where $\rho_i^{(2)}(\boldsymbol{r})$ is the frozen subsystem 2. For embedding calculations, i = FDE(s)(6) obtained with TF9W, whereas for the potential reconstruction, i = rec(k) with k the corresponding number of freeze-and-thaw cycles.

i	$\rho_i^{(2)} = \rho_i^{\mathrm{Li}^+}$	$\rho_i^{(2)} = \rho_i^{\mathrm{H}^-}$	$\rho_i^{(2)} = \rho_i^{\mathrm{Li}^+}$	$\rho_i^{(2)} = \rho_i^{\mathrm{H}^-}$
s.o.f.	4.35E-03	2.95E-01	4.35E-03	2.95E-01
TF9W	1.10E-03	1.62E-01	1.10E-03	1.62E-01
		starting: s.o.f.	starting: TF9W	
rec(0)	3.01E-03	2.12E-02	4.22E-05	1.52E-02
rec(1)	2.74E-03	1.69E-02	1.26E-09	1.35E-02
rec(2)	2.63E-03	1.62E-02	3.30E-05	1.33E-02
rec(3)	2.56E-03	1.59E-02	2.42E-04	1.13E-02
rec(4)	2.51E-03	1.57E-02	5.62E-04	7.50E-03
rec(5)	2.48E-03	1.56E-02	8.56E-04	9.44E-03
rec(6)	2.46E-03	1.55E-02	1.05E-03	1.14E-02
rec(7)	2.45E-03	1.54E-02	1.20E-03	1.25E-02
rec(8)	2.43E-03	1.53E-02	1.33E-03	1.31E-02
rec(9)	2.42E-03	1.53E-02	1.45E-03	1.36E-02
rec(10)	2.42E-03	1.52E-02	1.55E-03	1.39E-02

Figure 5.1 shows the deformation densities for embedding calculation with the TF9W functional, and potential reconstruction calculations with one and ten freeze-and-thaw cycles. Potential reconstruction brings a significant improvement upon TF9W. The difference between the rec(1) and rec(10) is not very apparent from the deformation density plot, and they both are in good agreement with the reference density. However, as discussed above, after rec(4), in fact, the quality of the reconstructed density continuously drops.

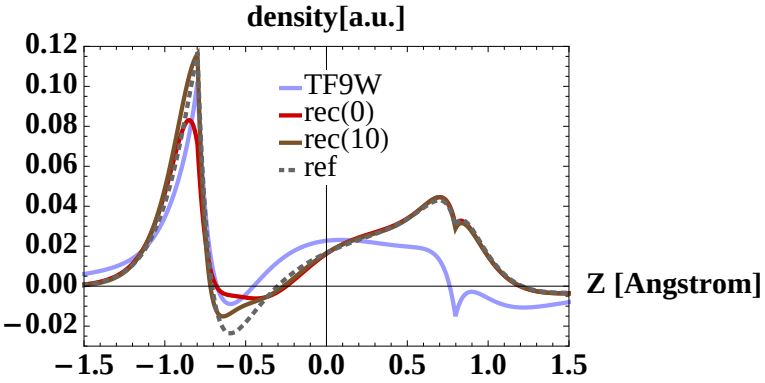


Figure 5.1: LiH: Deformation densities for embedding with the TF9W functional, and potential reconstruction calculations in a.u. Ref denotes $\rho^{\mathrm{ref}} - \rho^{(\mathrm{H}^-)} - \rho^{(\mathrm{Li}^+)}$, where ρ^{ref}, $\rho^{(\mathrm{H}^-)}$ and $\rho^{(\mathrm{Li}^+)}$ are the electron densities obtained from KS-DFT calculations on the full system, isolated H^- and Li^+ fragments; respectively. TF9W, rec(1), and rec(10) denote $\rho_i^{(\mathrm{H}^-)} + \rho_i^{(\mathrm{Li}^+)} - \rho^{(\mathrm{H}^-)} - \rho^{(\mathrm{Li}^+)}$, where $i =$ FDE(s)(6) and $\mathrm{rec(k)}$, with k the corresponding number of freeze-and-thaw cycles; respectively. The H^- atom is situated at $z = 0.80$ Å, Li^+ at $z = -0.80$ Å.

Assessment of the Potentials The effective potential of the subsystem H^- can be calculated exactly using the below series of equations.

$$h\psi(\boldsymbol{r}) = \varepsilon\psi(\boldsymbol{r}) \tag{5.8}$$

$$\Big(-\frac{\nabla^2}{2m} + \mathrm{V}(\boldsymbol{r})\Big)\psi(\boldsymbol{r}) = \varepsilon\psi(\boldsymbol{r}) \tag{5.9}$$

$$\mathrm{V}(\boldsymbol{r})\psi(\boldsymbol{r}) = \Big(\frac{\nabla^2}{2\mathrm{m}} + \varepsilon\Big)\psi(\boldsymbol{r}) \tag{5.10}$$

$$\mathrm{V}(\boldsymbol{r}) = \frac{\nabla^2\psi(\boldsymbol{r})}{2\mathrm{m}\psi(\boldsymbol{r})} + \varepsilon \tag{5.11}$$

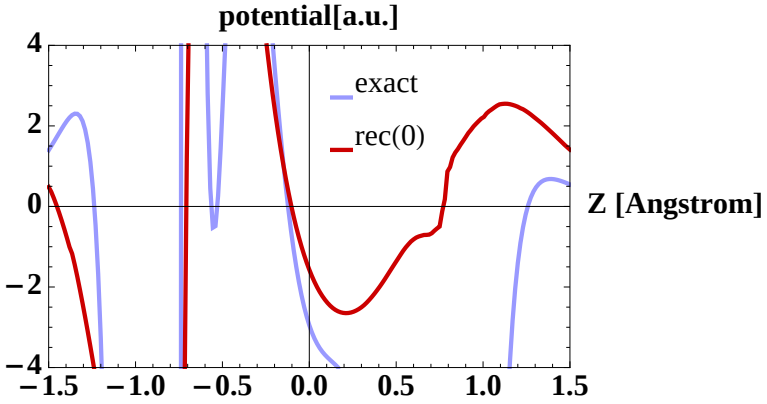
$$\psi(\boldsymbol{r}) = \sqrt{\rho(\boldsymbol{r})} \tag{5.12}$$

$$\mathrm{V}(\boldsymbol{r}) = \frac{\nabla^2\rho^{\frac{1}{2}}(\boldsymbol{r})}{2\mathrm{m}\rho^{\frac{1}{2}}(\boldsymbol{r})} + \varepsilon \tag{5.13}$$

$$\mathrm{V}(\boldsymbol{r}) = \frac{1}{4\mathrm{m}}\frac{\nabla^2\rho(\boldsymbol{r})}{\rho(\boldsymbol{r})} - \frac{1}{8\mathrm{m}}\frac{\nabla\rho(\boldsymbol{r})\nabla\rho(\boldsymbol{r})}{\rho^2(\boldsymbol{r})} + \varepsilon \tag{5.14}$$

It is the relation given in Eq. (5.12), which enables the calculation of the exact potential for a 2-electron system such as H^-, where the density obtained from the potential reconstruction calculation is employed. This exact potential $\mathrm{V}(\boldsymbol{r})$ can then be compared with the reconstructed potential. To be able to make a more direct comparison, we subtracted the initial potential $v_{\mathrm{initial}}(\boldsymbol{r})$ from $\mathrm{V}(\boldsymbol{r})$, which can be analysed against only the reconstructed part of the final potential, i.e. $\sum_t b_t g_t(\boldsymbol{r})$. Furthermore, we calculated $\mathrm{V}(\boldsymbol{r})$ using our least accurate density obtained from rec(0) starting from s.o.f. densities, as well as our most accurate density obtained from rec(1) starting from TF9W densities. This comparison is shown in Figure 5.2. Whether starting from s.o.f. or TF9W densities, the reconstructed potentials capture the qualitative trend, however, do not result in the correct magnitude. Although, the reconstructed potential started from s.o.f. densities does not have the correct behaviour nearby the nucleus of the frozen subsystem Li^+ (at $z = -0.50\,\text{Å}$), where the potential becomes attractive. It is noteworthy that, this attractive behaviour is not present in the exact potential calculated using the electron density resulted from the reconstructed potential started from TF9W densities.

(a)



(b)

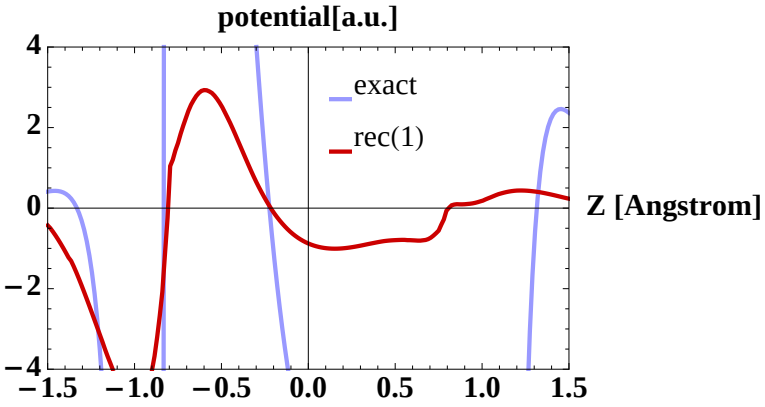
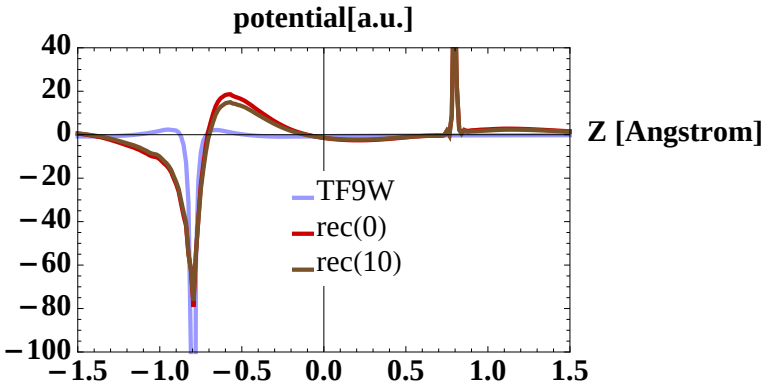


Figure 5.2: LiH: Comparison of exact and reconstructed potentials using our (a) least accurate electron density obtained by starting the potential reconstruction from isolated fragments densities and (b) most accurate electron density obtained by starting the potential reconstruction from FDE(s)(6) densities with TF9W rec(0) and rec(1) denote $\sum_t b_t g_t(\boldsymbol{r})$ with no and one freeze-and-thaw calculation; respectively and exact refers to $V_{\boldsymbol{r}} - v^{\text{initial}}(\boldsymbol{r})$. The H^- atom is situated at $z = 0.80$ Å, Li^+ at $z = -0.80$ Å.

Figure 5.3 shows a comparison of embedding potentials and nonadditive kinetic-energy components obtained from rec(0) and rec(10) of potential reconstruction calculations started from s.o.f. densities, with those obtained from an embedding calculation with TF9W. rec(0) has, therefore, the same input density as Tf9W, and allows a direct comparison. Potential reconstruction brings an improvement upon the embedding with TF9W. rec(0) results in a $v_{\mathrm{T}}^{\mathrm{nadd}}[\rho^{(1)}, \rho^{(2)}]$ potential, which is more repulsive around the nucleus of the frozen subsystem Li^{+} than that obtained with TF9W. This is in a better agreement with the correct behaviour as the repulsive kinetic-energy component compensates to a greater degree the nuclear attraction. Similarly, the embedding potential around the frozen nucleus (Li^{+}) is far less attractive than that of TF9W. Furthermore, TF9W completely misses the repulsive peak at the nucleus of the active subsystem H^{-}.

Nonadditive kinetic-energy component obtained from rec(0) and rec(10) differ in the region nearby the frozen nucleus (at $z = -0.50$ Å), where that of rec(0) is slightly more repulsive than that of rec(10). Whereas, for the embedding potentials, no significant difference is apparent.

(a)



(b)

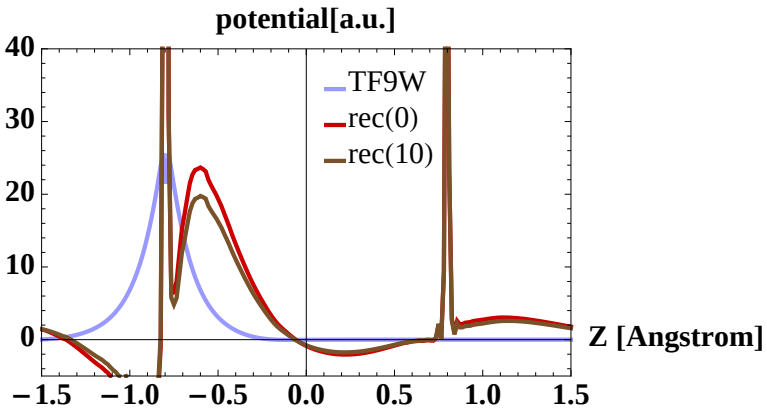


Figure 5.3: LiH: (a) Embedding and (b) Nonadditive kinetic-energy potentials obtained from the embedding with the TF9W functional and potential reconstruction calculation with one freeze-and-thaw calculation denoted as rec(1). The H^- atom is situated at $z = 0.80$ Å, Li^+ at $z = -0.80$ Å.

5.3.2 Li_2

Li_2 bond length is 2.68 Åand the Li^- atom is situated at $z = 1.34$ Å, Li^+ at $z = -1.34$ Å, where z is the bond axis. In all the plots Li^- is the active subsystem.

Assessment of the Density The effect of freeze-and-thaw cycles on the the dipole moment difference for the total system $\Delta\mu$ (Debye), and the integrated difference density Δ_{abs} as well as on the negative density is shown in Tables 5.3 and 5.4; respectively. As expected, embedding with TF9W significantly improves upon s.o.f., resulting in a higher quality of density with smaller negative density for both subsystems.

Potential reconstruction started from s.o.f. results in a similar picture as observed for the LiH molecule. It improves the quality of the electron density upon TF9W as well as s.o.f.. This improvement is consistently present in successive freeze-and-thaw cycles. However, similarly, from rec(8) to rec(10), the difference between successive cycles gets smaller (4.0E-4) as shown in Table 5.3. This trend is also present in the negative densities (see Table 5.4).

On the other hand, potential reconstruction scheme starting from TF9W densities results in a higher quality of the electron density, with respect to that of starting from s.o.f., already in rec(0), as observed also for the LiH molecule. However, for Li_2, the quality of the density as well as the negative density continuously decrease from rec(2) to rec(10). Remarkably, the electron density of the Li^- fragment obtained from rec(1), $\rho^{(\text{Li}^-)}_{\text{rec}(1)}(\boldsymbol{r})$ is positive everywhere in space. This is the fulfilment of the first condition described in Sec. 5.2.1. Given that $\rho^{(\text{Li}^-)}_{\text{rec}(1)}(\boldsymbol{r})$ is v_s-representable (the second condition), the next freeze-and-thaw cycle, i.e. rec(2), should in principle result in the exact density $\rho^{(\text{Li}^-)}_{\text{rec}(2)}(\boldsymbol{r})$ density, such that $\rho^{(\text{tot})}(\boldsymbol{r}) = \rho^{(\text{Li}^-)}_{\text{rec}(2)}(\boldsymbol{r}) + \rho^{(\text{Li}^+)}_{\text{rec}(2)}(\boldsymbol{r})$. However, rec(2) does not improve upon the result of rec(1) but instead worsens it.

Table 5.3: Li_2: Effect of freeze-and-thaw cycles on the the dipole moment difference for the total system $\Delta\mu$ (Debye), and the integrated difference density Δ_{abs} (Eq. (5.5)). s.o.f. denotes the sum of isolated fragments. Potential reconstruction calculations were started from the sum of isolated fragments as well as FDE(s)(6) densities obtained with TF9W.

	Δ_{abs}	$\Delta\mu$	Δ_{abs}	$\Delta\mu$
s.o.f.	1.7098	12.839	0.6321	12.839
TF9W	0.8736	4.2048	0.6409	4.2048
	starting: s.o.f.		starting: TF9W	
rec(0)	0.0968	0.0890	0.0558	0.1112
rec(1)	0.0943	0.0918	0.0505	0.0669
rec(2)	0.0927	0.0928	0.0564	0.0964
rec(3)	0.0903	0.0963	0.0606	0.1181
rec(4)	0.0893	0.1000	0.0632	0.1267
rec(5)	0.0887	0.1021	0.0653	0.1310
rec(6)	0.0882	0.1036	0.0674	0.1334
rec(7)	0.0877	0.1040	0.0694	0.1321
rec(8)	0.0873	0.1047	0.0712	0.1299
rec(9)	0.0869	0.1051	0.0726	0.1270
rec(10)	0.0865	0.1054	0.0740	0.1246

Table 5.4: Li_2 Effect of freeze-and-thaw cycles on the negative density, i.e. integrated negative part of the difference $\rho^{(\mathrm{tot})}(\boldsymbol{r}) - \rho_i^{(2)}(\boldsymbol{r})$ where $\rho_i^{(2)}(\boldsymbol{r})$ is the frozen subsystem 2. For embedding calculations, i = FDE(s)(6) obtained withTF9W, whereas for the potential reconstruction, i = rec(k) with k the corresponding number of freeze-and-thaw cycles.

i	$\rho_i^{(2)} = \rho_i^{\mathrm{Li}^+}$	$\rho_i^{(2)} = \rho_i^{\mathrm{Li}^-}$	$\rho_i^{(2)} = \rho_i^{\mathrm{Li}^+}$	$\rho_i^{(2)} = \rho_i^{\mathrm{Li}^-}$
s.o.f.	1.17E-03	8.52E-01	1.17E-03	8.52E-01
TF9W	5.15E-04	2.50E-01	5.15E-04	2.50E-01
		starting: s.o.f.	starting: TF9W	
rec(0)	1.04E-03	4.09E-02	4.04E-06	2.10E-02
rec(1)	9.75E-04	3.97E-02	0	2.05E-02
rec(2)	8.34E-04	3.91E-02	2.29E-06	2.38E-02
rec(3)	7.69E-04	3.81E-02	1.93E-06	2.56E-02
rec(4)	7.26E-04	3.76E-02	5.30E-05	2.65E-02
rec(5)	6.94E-04	3.73E-02	9.76E-05	2.73E-02
rec(6)	6.73E-04	3.71E-02	1.43E-04	2.82E-02
rec(7)	6.53E-04	3.69E-02	1.85E-04	2.90E-02
rec(8)	6.35E-04	3.67E-02	2.29E-04	2.97E-02
rec(9)	6.19E-04	3.65E-02	2.68E-04	3.04E-02
rec(10)	6.04E-04	3.63E-02	3.00E-04	3.10E-02

In order to see whether the electron density $\rho^{(\mathrm{Li}^-)}_{\mathrm{rec}(1)}(\boldsymbol{r})$ is v_s-representable, we plotted the electron density difference $\rho^{\mathrm{ref}} - \rho^{(\mathrm{Li}^+)}_{\mathrm{rec}(1)}$ along the bond axis where ρ^{ref} is the electron density obtained from a KS-DFT calculation on the full system, and $\rho^{(\mathrm{Li}^-)}_{\mathrm{rec}(1)}$ is the electron density of the subsystem Li^- obtained from the potential reconstruction calculation with one freeze-and-thaw cycle. This is shown in Figure 5.4. $\rho^{\mathrm{ref}} - \rho^{(\mathrm{Li}^+)}_{\mathrm{rec}(1)}$ appears to be zero around $z = -1$ Å, and therefore, is not v_s-representable. As described in Section 5.2.1, at this surface where $\rho^{\mathrm{ref}} - \rho^{(\mathrm{Li}^+)}_{\mathrm{rec}(1)}$ is zero, the potential would have to be singular [$v_{\mathrm{T}}^{\mathrm{nadd}}(\boldsymbol{r}) = \infty$], which is very difficult to represent in practical calculations. Thus, even though the first condition is fulfilled, since the the second condition is not met, the exact density cannot be achieved

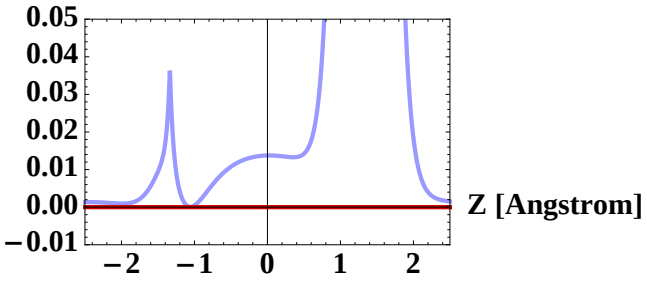
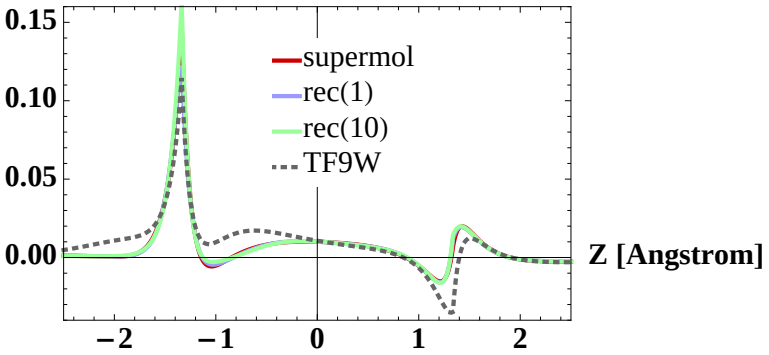


Figure 5.4: Li_2: $\rho^{\mathrm{ref}} - \rho^{(\mathrm{Li}^+)}_{\mathrm{rec}(1)}$ along the bond axis where ρ^{ref} is the electron density obtained from a KS-DFT calculation on the full system, and $\rho^{(\mathrm{Li}^-)}_{\mathrm{rec}(2)}$ is the electron density of the subsystem Li^- obtained from the potential reconstruction calculation with one freeze-and-thaw cycles. The Li^- atom is situated at $z = 1.34$ Å, Li^+ at $z = -1.34$ Å.

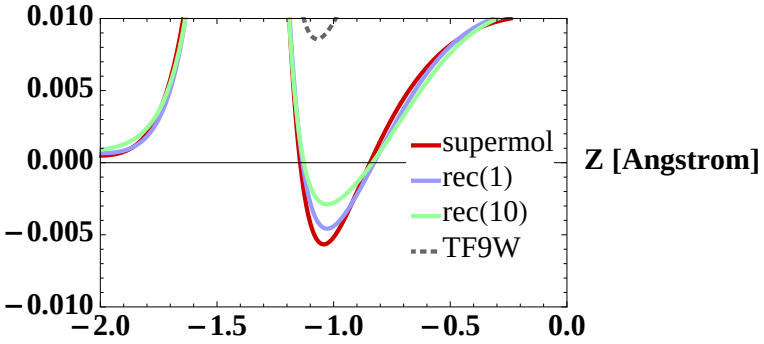
Figure 5.5 shows the deformation densities for embedding calculation with the TF9W functional, and rec(1) and rec(10) of potential reconstruction calculations started from TF9W densities. Potential reconstruction brings a significant improvement upon TF9W. The difference between the rec(1) and rec(10) is not very apparent from the deformation density plot, and they both are in good agreement with the reference density. However, as discussed above, after the first freeze-and-thaw cycle, the quality of the reconstructed density in fact drops. Deformation density obtained from the potential reconstruction started from s.o.f. densities has a very similar picture and, therefore, is not shown separately. Figures 5.5 (b) and (c), which zoom in the regions near the nuclei of Li^+ and Li^-; respectively, highlight the differences between rec(1) and rec(10).

Reconstructed embedding potential and the kinetic-energy component are not shown for the Li_2 molecule, since it has a similar picture to that of LiH.

(a)



(b)



(c)

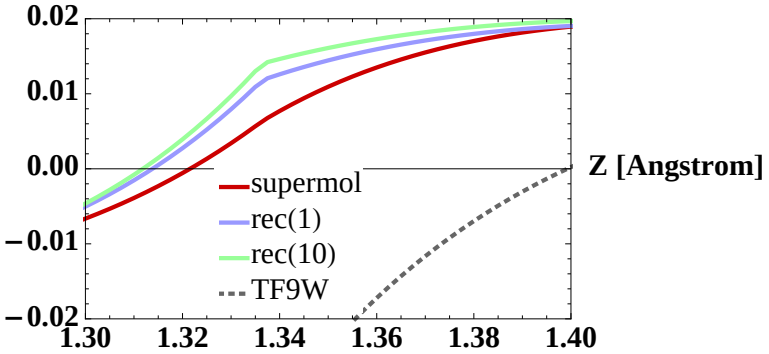


Figure 5.5: Li_2: Deformation densities for embedding with the TF9W functional, and potential reconstruction calculations in a.u. Supermole denotes $\rho^{\mathrm{ref}} - \rho^{(\mathrm{Li}^-)} - \rho^{(\mathrm{Li}^+)}$, where ρ^{ref}, $\rho^{(\mathrm{Li}^-)}$ and $\rho^{(\mathrm{Li}^+)}$ are the electron densities obtained from KS-DFT calculations on the full system, isolated Li^- and Li^+ fragments; respectively. TF9W, rec(1), and rec(10) denote $\rho_i^{(\mathrm{Li}^-)} + \rho_i^{(\mathrm{Li}^+)} - \rho^{(\mathrm{Li}^-)} - \rho^{(\mathrm{Li}^+)}$, where $i =$ FDE(s)(6) and $\mathrm{rec(k)}$, with $k = 1, 10$ the number of freeze-and-thaw cycles; respectively. The Li^- atom is situated at $z = 1.34$ Å, Li^+ at $z = -1.34$ Å.

5.4 Conclusions

We used the potential reconstruction scheme presented in 5.1, and built up on the work of Fux *et al.*[86] by investigating the validity of the freeze-and-thaw scheme in conjunction with the potential reconstruction. We analysed the effect of freeze-and-thaw cycles, the necessity of which was investigated earlier,[69] on the quality of reconstructed embedding potentials and electron densities. We used two model systems with covalent bonding: LiH and Li_2, which we partitioned into Li^+ and H^-; Li^+ and Li^- subsystems; respectively. The advantage of such a partitioning is that it reduces the negativity of $\rho^{(1)}(\boldsymbol{r})$ and, hence, is a great attempt to fulfil the first condition described above. Furthermore, it provides two strongly different starting points for these simple model systems for which one can analyse the effect of freeze-and-thaw cycles.

We started the potential reconstruction calculations from the the superposition of the densities obtained by KS calculations on the isolated fragments (s.o.f.) as well as from FDE(s)(6) densities obtained with TF9W. To assess the quality of the reconstructed electron densities, we used the dipole moment difference for the total system $\Delta\mu$ (Eq. (4.15)), and the integrated difference density, Δ_{abs} (Eq. (5.5)) that is also used for choosing λ within the potential reconstruction scheme. We also investigated the effect of freeze-and-thaw cycles on the negative density as well as deformation density.

For both LiH and Li_2, potential reconstruction scheme started with the s.o.f. densities improves the quality of the electron density upon TF9W as well as s.o.f.. This improvement is consistently present in successive freeze-and-thaw cycles. However, from rec(8) to rec(10), the difference between successive cycles gets smaller (1.0E-4 and 4.0E-04 for LiH and Li_2; respectively). This trend is also present in the negative densities.

On the other hand, we showed that for both LiH and Li_2, potential reconstruction scheme starting from TF9W densities results in a higher quality of the electron density, with respect to that of starting from s.o.f., already in rec(0). In the case of the LiH molecule, it is noteworthy that these TF9W starting densities have a lower quality than that of s.of. but a lower negative density for both subsystems. The improvement obtained, therefore, shows the importance of negative density in the starting point.

For LiH, the freeze-and-thaw cycles result in an improvement until the fifth cycle, then the quality of reconstructed densities continuously decreases. For Li_2, already after the first cycle, the improvement stops. Remarkably, the electron

density of the Li^- fragment obtained from this first cycle, is positive everywhere in space. This is the fulfilment of the first condition described in Sec. 5.2.1. Given that the electron density is v_s-representable (the second condition), the next freeze-and-thaw cycle, i.e. the second cycle, should in principle result in the exact density. However, the second cycle does not improve upon the result of the first one but instead worsens it. The reason for this phenomena certainly deserves attention and is subject to future research. Identifying the causes for this problem can help improve the current reconstruction scheme.

Furthermore, for LiH, given that the system is v_s-representable, we calculated the exact potential for the 2-electron H^- fragment, which we compared with that obtained from the reconstruction scheme. We found that reconstructed potentials capture the qualitative trend, however, fail to obtain the correct magnitude.

Reconstructed embedding potentials and the kinetic-energy component improves upon TF9W. Expectedly, compared to TF9W, reconstructed potentials are far less attractive around the frozen nucleus since the nonadditive kinetic-energy component compensates for the electrostatic attraction. Furthermore, we found that TF9W completely misses a repulsive peak at the nucleus of the active subsystem.

Chapter 6

Subsystem Density Functional Theory with Bond Energy Decomposition Analysis [1]

Abstract

In this chapter, an explicit expression for the dispersion (D) energy in conjunction with Kohn-Sham (KS) density functional theory and frozen-density embedding (FDE) to calculate interaction energies between DNA base pairs and a selected set of amino-acid pairs in the hydrophobic core of a small protein Rubredoxin (Rd) is employed. This data to assess the accuracy of an FDE-D approach for the calculation of intermolecular interactions.

To better analyze the calculated interaction energies, a new energy decomposition scheme is proposed that is similar to the well-known KS bond formation analysis [F. M. Bickelhaupt and E. J. Baerends, Rev. Comput. Chem. **15**, 1 (2000)], but differs in the densities used to define the bond energy. The individual subsystem densities of the FDE approach sum to the total energy which makes it possible to define bond energies in terms of promotion energies and an explicit interaction energy. For the systems considered, only a few freeze-and-thaw cycles suffice to reach convergence in these individual bond energy components, illustrating the potential of FDE-D as an efficient method to calculate intermolecular

[1] The contents of this chapter will be published as S. M. Beyhan, A. W. Götz, and L. Visscher, to be submitted

interactions.

6.1 Introduction

Intermolecular interactions play a pivotal role in structure formation of biomolecular systems and govern processes like protein folding, molecular recognition and stacking of nucleobases.[148–150] With the ever-increasing computational resources, explicit study of the corresponding interaction energies by means of quantum mechanical calculations comes within reach. Calculation of interaction energies for biological systems with highly accurate *ab initio* methods is, however, only possible for benchmark purposes[151–155] as routine application is still hampered by the steep scaling of computational costs with system size. Density Functional Theory (DFT)[156] scales better and is very successful in treating H-bonding but the available approximations to the exchange-correlation (XC) functional typically describe dispersion interactions rather poorly.[48–51]

One possibility to improve this behavior is to move towards truly nonlocal density functionals (vdW-DF[157,158,158]), but most approaches rather combine a standard XC functional with an additional treatment of dispersion interactions (for recent papers see Refs. 52–55). An elegant way to do so[159] is to use symmetry-adapted perturbation theory (SAPT)[160] in which a DFT description of monomers is combined with an explicit treatment of intermolecular interactions. More approximate, but popular due to their ease of implementation and computational efficiency, is the class of DFT-D methods in which an approximate expression based on the limiting C_6R^{-6} term of the dispersion interaction is used.[161–166] DFT-D is nowadays the most widely used approach to the dispersion problem in biomolecular systems and various DFT-D implementations have been reported,[161–163,167–173] following pioneering work of in particular Grimme.[162,167,174] Examples of successful treatments of biologically relevant systems can e.g. be found in Refs. 52, 175.

Frozen Density Embedding and Interaction Energy Even though DFT calculations are much cheaper than correlated wave-function based approaches, it is still very expensive to carry out supermolecular calculations for biological systems. A convenient approach for modelling biological systems is therefore the use of subsystem methods in which only a region of interest is treated at a high level of theory and the environment is dealt with at a lower level of theory.[3–6] The frozen-density method, proposed by Wesolowski and Warshel,[59] is particularly attractive because this subsystem theory is a reformulation of DFT that is in principle exact. The FDE approach is often used to calculate molecular properties of solvated systems[62,65,89] or, in its generalization to time-dependent DFT,[90,93] to describe

local electronic excitations and couplings between such excitations.[63,66,99,100] In the present context we are interested in the calculation of interaction energies of weakly interacting systems in which the method is known to perform rather well,[101–103].SMB2

FDE is a subsystem formulation[67] of DFT and in most cases the density is partitioned into two subdensities that each correspond to an integer number of electrons. The total electron density $\rho_{\mathrm{tot}}(\boldsymbol{r})$ is then given by

$$\rho^{(\mathrm{tot})}(\boldsymbol{r}) = \rho^{(1)}(\boldsymbol{r}) + \rho^{(2)}(\boldsymbol{r}). \tag{6.1}$$

The DFT energy for this two-partitioning is written as

$$E[\rho^{(1)}, \rho^{(2)}] = \sum_{i=1}^{2} E^{(i)}[\rho^{(i)}] + E_{\mathrm{int}}[\rho^{(1)}, \rho^{(2)}], \tag{6.2}$$

where $E^{(i)}[\rho^{(i)}]$ is the KS-DFT energy of subsystem i:

$$\begin{aligned} E^{(i)}[\rho^{(i)}] &= T_s[\rho^{(i)}] + \int v^{(i)}_{\mathrm{nuc}}(\boldsymbol{r})\rho^{(i)}(\boldsymbol{r})\mathrm{d}\boldsymbol{r} \\ &\quad \frac{1}{2}\int \frac{\rho^{(i)}(\boldsymbol{r})\rho^{(i)}(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r'}|}\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{r'} + \mathrm{E_{xc}}[\rho^{(\mathrm{i})}]. \end{aligned} \tag{6.3}$$

Here, $v^{(i)}_{\mathrm{nuc}}(\boldsymbol{r})$ is the electrostatic potential due to the nuclei in subsystem i and $E_{\mathrm{xc}}[\rho^{(i)}]$ is the XC energy of subsystem i. The interaction energy E_{int} between the two subsystems is

$$\begin{aligned} E_{\mathrm{int}}[\rho^{(1)}, \rho^{(2)}] &= \sum_{i\neq j}^{2} \int v^{(i)}_{\mathrm{nuc}}(\boldsymbol{r})\rho^{(j)}(\boldsymbol{r})\mathrm{d}\boldsymbol{r} + \int \frac{\rho^{(1)}(\boldsymbol{r})\rho^{(2)}(\boldsymbol{r})}{|\boldsymbol{r}-\boldsymbol{r'}|}\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{r'} \\ &\quad + T_s^{\mathrm{nad}}[\rho^{(1)}, \rho^{(2)}] + E_{\mathrm{xc}}^{\mathrm{nad}}[\rho^{(1)}, \rho^{(2)}] \end{aligned} \tag{6.4}$$

where the nonadditive contribution to the kinetic energy $T_s^{\mathrm{nad}}[\rho^{(1)}, \rho^{(2)}]$ and the nonadditive contribution to the XC energy $E_{\mathrm{xc}}^{\mathrm{nad}}[\rho^{(1)}, \rho^{(2)}]$ are given by

$$T_s^{\mathrm{nad}}[\rho^{(1)}, \rho^{(2)}] = T_s[\rho^{\mathrm{tot}}] - \sum_{i}^{2} T_s[\rho^{(i)}] \tag{6.5}$$

and

$$E_{\mathrm{xc}}^{\mathrm{nad}}[\rho^{(1)}, \rho^{(2)}] = E_{\mathrm{xc}}[\rho^{\mathrm{tot}}] - \sum_{i}^{2} E_{\mathrm{xc}}[\rho^{(i)}]. \tag{6.6}$$

These terms can be evaluated once explicit density functionals for the XC energy, $E_{\rm xc}[\rho]$, as well as for the kinetic energy (KE) of the non-interacting reference KS system, $T_{\rm s}[\rho]$ are chosen. The latter is only used for the contribution to the interaction energy, since the kinetic energy of the subsystems can be evaluated from the KS orbitals. To obtain the orbitals of the subsystems, a set of coupled KS-like equations is derived by minimizing the total energy functional in Eq. 6.2 with respect to the electron density $\rho^{(1)}$ of the first subsystem (1) while keeping the electron density of the other subsystem (2) frozen. The resulting equation reads

$$\Big[-\frac{\nabla^2}{2} + v_{\rm eff}^{\rm KSCED}[\rho^{(1)},\rho^{(2)}](\boldsymbol{r})\Big]\phi^{(1)}(\boldsymbol{r}) = \epsilon_i \phi^{(1)}(\boldsymbol{r}), \tag{6.7}$$

with the effective potential given by

$$v_{\rm eff}^{\rm KSCED}[\rho^{(1)},\rho^{(2)}](\boldsymbol{r}) = v_{\rm eff}^{\rm KS}[\rho^{(1)}](\boldsymbol{r}) + v_{\rm eff}^{\rm emb}[\rho^{(1)},\rho^{(2)}](\boldsymbol{r}), \tag{6.8}$$

where $v_{\rm eff}^{\rm KS}[\rho^{(1)}](\boldsymbol{r})$ is the usual KS effective potential of subsystem 1

$$\begin{aligned} v_{\rm eff}^{\rm KS}[\rho^{(1)}](\boldsymbol{r}) &= v_{\rm nuc}^{(1)}(\boldsymbol{r}) + \int \frac{\rho^{(1)}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r}' \\ &\quad + \frac{\delta E_{\rm xc}[\rho]}{\delta\rho}\Big|_{\rho=\rho^{(1)}(\boldsymbol{r})}. \end{aligned} \tag{6.9}$$

containing, respectively, the nuclear potential, the Coulomb potential and the XC potential evaluated with only the density of subsystem 1. The effect of subsystem 2 is described by the embedding potential $v_{\rm eff}$ that depends on both densities and reads

$$\begin{aligned} v_{\rm eff}^{\rm emb}[\rho^{(1)},\rho^{(2)}](\boldsymbol{r}) &= v_{\rm nuc}^{(2)}(\boldsymbol{r}) + \int \frac{\rho^{(2)}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r}' \\ &\quad + \frac{\delta E_{\rm xc}[\rho]}{\delta\rho}\Big|_{\rho=\rho^{(\rm tot)}(\boldsymbol{r})} - \frac{\delta E_{\rm xc}[\rho]}{\delta\rho}\Big|_{\rho=\rho^{(1)}(\boldsymbol{r})} \\ &\quad + \frac{\delta T_{\rm s}[\rho]}{\delta\rho}\Big|_{\rho=\rho^{(\rm tot)}(\boldsymbol{r})} - \frac{\delta T_{\rm s}[\rho]}{\delta\rho}\Big|_{\rho=\rho^{(1)}(\boldsymbol{r})}, \end{aligned} \tag{6.10}$$

where $v_{\rm nuc}^{(2)}(\boldsymbol{r})$ denotes the external potential due to the nuclei of subsystem 2 and $\rho^{(\rm tot)}(\boldsymbol{r}) = \rho^{(1)}(\boldsymbol{r}) + \rho^{(2)}(\boldsymbol{r})$ is the electron density of the whole system.

If the KE density functional $T_s[\rho]$ were exact, supermolecular KS-DFT results could be reproduced exactly by the subsystem calculation. This requires

two additional conditions on the fixed density $\rho^{(2)}(\boldsymbol{r})$: the remaining density $\rho^{(1)} = \rho^{(\mathrm{tot})} - \rho^{(2)}$ should be non-negative everywhere in space and this density should be non-interacting v_{s}-representable so that it can be obtained via KS-DFT.[60,82] Since the first condition is difficult to fulfill exactly by simple trial densities, often a so-called "freeze-and-thaw" (FT) procedure[68] is introduced in which the subsystem densities are adjusted in an iterative fashion. More importantly, however, are deficiencies in the best known approximations to the exact KE functional. These are particularly apparent for ligand-metal bonds in coordination complexes[SMB2, SMB5, 77, 83] and better approximations need be developed in order to allow for quantitative accuracy. On the other hand, the FDE scheme is able to handle even strong hydrogen bonds[82] and can also describe weakly interacting systems well.[78, 81, 104, 176] Due to error cancellations, FDE can sometimes appear to even outperform KS-DFT with an appropriate combination of functionals for E_{XC} and T_s. Often used choices are to employ local density approximation (LDA) for both the XC[177] and the KE[178, 179] functional or to use the PW91 generalized gradient approximation (GGA)[180] functional in combination with the PW91K KE functional.[71, 181] The latter combination works well for π–stacked base pairs.[176]

In an earlier study, we computed the interaction energies of a wide range of systems to further assess the performance of various kinetic energy functionals.[SMB2] Given the satisfactory reproduction of KS data for typically biomolecular interactions we now aim to add dispersion corrections and move towards a quantitative description of interaction energies in non-covalently bound systems. In our view, this new FDE-D scheme should be regarded as a computationally efficient approximation to KS-DFT, reproducing in the ideal case KS-DFT-D. The dispersion correction is in this context viewed as a correction to the XC interaction energy that has no influence on the error in the kinetic-energy part.

Another goal of the current work is to introduce an energy decomposition scheme that can provide more insight in trends in the interaction energies. This is inspired by the well-known Morokuma scheme[182] that is used to analyze bond formation in KS theory[28, 183] and subsystems embedded in and interacting with a larger environment (e.g. protein, solvent).[184–189] In this analysis, given the KS orbitals and electron densities $\rho^{(1)}_{\mathrm{frag}}$ and $\rho^{(2)}_{\mathrm{frag}}$ of the two isolated fragments at the dimer geometry, the bond energy E^{bond} is defined as

$$E_{\mathrm{bond}} = \Delta E_{\mathrm{elstat}} + \Delta E_{\mathrm{Pauli}} + \Delta E_{\mathrm{oi}}, \tag{6.11}$$

where $\Delta E_{\mathrm{elstat}}$ is the electrostatic interaction calculated using the unperturbed densities of the fragments, $\Delta E_{\mathrm{Pauli}}$ is the energy change caused by orthogonalization of fragment wave functions, and ΔE_{oi} arises from the recombination of the orthogonalized fragment orbitals to form the supermolecular KS wave func-

tion. The repulsive Pauli contribution can be further decomposed into a change in kinetic energy (ΔT^0) and into a change of potential energy $\Delta V_{\mathrm{Pauli}}$:

$$\Delta E_{\mathrm{Pauli}} = \Delta T^0 + \Delta V_{\mathrm{Pauli}}. \tag{6.12}$$

In this KS bond formation analysis, the orthogonalization of the fragment wave functions thus plays a crucial role. This orthogonalization pushes density out of the bonding region towards the nuclei which raises the kinetic energy but lowers the potential energy. The net effect is a strongly positive energy that can be used to quantify the steric repulsion between the fragments.

Within FDE theory fragment orbitals remain non-orthogonal and provide subsystem densities that sum to the total density (Eq. 6.1). In order to analyze the bond energy we may thus keep the $\Delta E_{\mathrm{elstat}}$ term, but need to replace the $\Delta E_{\mathrm{Pauli}}$ and ΔE_{oi} terms that are based on the orthogonalized orbitals. We propose to rewrite the bond energy E_{bond} as:

$$\Delta E_{\mathrm{bond}} = \sum_{i=1}^{2} E^{(i)}_{\mathrm{prom}}[\rho^{(i)}] + E_{\mathrm{int}}[\rho^{(1)}, \rho^{(2)}], \tag{6.13}$$

where $E_{\mathrm{int}}[\rho^{(1)}, \rho^{(2)}]$ is the interaction energy between the two subsystems as given by Eq. 6.4 and $E^{(i)}_{\mathrm{prom}}[\rho^{(i)}]$ is the so-called *promotion energy* of subsystem i:

$$E^{(i)}_{\mathrm{prom}} = \Delta E^{(i)}_{\mathrm{Hartree}} + \Delta E^{(i)}_{\mathrm{xc}} + \Delta T^{(i)}_{s}, \tag{6.14}$$

where $\Delta E^{(i)}_{\mathrm{xc}}$, $\Delta T^{(i)}_{s}$ and $\Delta E^{(i)}_{\mathrm{Hartree}}$ refer to the differences in XC, kinetic and Hartree energies induced by the change from the fragment $\rho^{(i)}_{\mathrm{frag}}$ density to the fully interacting subsystem density $\rho^{(i)}$. Due to the non-linear dependence on the density, the XC and KE terms need to be calculated as a difference:

$$\Delta E^{(i)}_{\mathrm{xc}}[\rho^{(i)}] = E_{\mathrm{xc}}[\rho^{(i)}] - E_{\mathrm{xc}}[\rho^{(i)}_{\mathrm{frag}}], \tag{6.15}$$

$$\Delta T^{(i)}_{s}[\rho^{(i)}] = T_{s}[\rho^{(i)}] - T_{s}[\rho^{(i)}_{\mathrm{frag}}]. \tag{6.16}$$

The Hartree term (in which the nuclear attraction and electron repulsion terms are summed) can also be calculated using the density difference $\Delta\rho^{(i)} = \rho^{(i)} -$

$\rho^{(i)}_{\mathrm{frag}}$, which improves numerical accuracy if a numerical integration scheme is employed for its calculation:

$$\begin{aligned}\Delta E^{(i)}_{\mathrm{Hartree}}[\rho^{(i)}] &= \int v^{(i)}_{\mathrm{nuc}}(\boldsymbol{r})\Delta\rho^{(i)}\mathrm{d}\boldsymbol{r} + \int \frac{\rho^{(\mathrm{i})}(\boldsymbol{r})\Delta\rho^{(\mathrm{i})}(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r'}|}\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{r'} \\ &\quad + \frac{1}{2}\int \frac{\Delta\rho^{(i)}(\boldsymbol{r})\Delta\rho^{(i)}(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r'}|}\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{r'}\end{aligned} \tag{6.17}$$

This decomposition of the bond energy does depend on the final partitioning of the total electron density into subsystem densities and in principle requires an additional criterium to make the partitioning unique. Because T_s^{nad} is responsible for the errors relative to the supermolecular DFT approach, a possible strategy is to chose a partitioning in which this contribution to the interaction energy vanishes. This is also the idea behind the exact embedding scheme proposed recently by Manby[190] in which localized orbitals of subsystems are kept orthogonal to each other by level-shifting techniques. One thereby eliminates the non-additive KE component of the interaction energy at the expense of increasing the promotion energies. Another choice is to minimize the promotion energies, and attribute the bond energy as much as possible to an explicit interaction energy. We suspect that the current procedure of applying FT cycles starting from the isolated fragment densities will typically give results close to this choice.

The analysis of the bond energy in terms of promotion and interaction energies can provide insight in the effect of the FT procedure which is typically interpreted as describing the environment polarization. In the applications described below we therefore calculate the changes in each individual component in each FT cycle. To test this approach, we chose molecules in which hydrogen bonds and dispersion interactions dominate. [2] Such systems can be readily treated with the current realizations of FDE. We selected in this study the DNA base pairs in the BP8/05 data set[191] of Truhlar, which we also used in our earlier work,[SMB2] and the Rubredoxin (Rd) protein for which accurate reference data are available from the work of Vondrášek *et al.*[192,193] In order to assess the efficiency of FDE, we will also discuss briefly the computational cost of the FDE and reference KS calculations.

The next section contains the computational details, in Section 6.3 we present our results where in Section 6.3.1 we discuss the convergence behaviour of total interaction energy contributions, in Section 6.3.2 we analyse the accuracy of FDE

[2]Note that we will use the term 'bond energy' to denote the sum of promotion and interaction energies even though the systems considered in this work are not always considered 'bonded' in conventional nomenclature

bond energies, and in Section 6.3.3 we discuss the computational cost of the FDE relative to reference KS calculations. Section 6.4 contains our conclusions.

6.2 Computational Details

All calculations were performed with the Amsterdam Density Functional (ADF) program.[194–196] We used the TZP basis set of the ADF basis set library which is a triple-ζ valence/double-ζ core all-electron Slater basis augmented with one set of polarization functions. The numerical integration parameters of ADF was set to an accuracy of 10 digits while for the self-consistent field (SCF) procedure the ADF default setting was used. We used the BLYP XC functional which is equivalent to the Becke[197] gradient correction and the Lee–Yang–Parr[198–200] correlation correction in both the KS-DFT as well as the FDE calculations. Since ADF uses Slater-type basis functions the Coulomb contribution to the interaction energy (Eq. 6.4) is calculated using a fitted density. To be able to correct for the fit error in this term, all FDE calculations were done using a full, supermolecular integration grid. More details on this implementation of the Coulomb energy for FDE can be found in our earlier work.[SMB2] The dispersion contribution was evaluated using the 2006 Grimme correction[167] for both KS-DFT and FDE calculations. This dispersion correction is added to the total bonding energy with the global scaling factor 1.20 for the BLYP XC functional. All the KS-DFT energies were corrected for the basis set superposition error (BSSE) with the counterpoise technique.[201] To setup and execute all our calculations and then to retrieve the data, we used PyADF[SMB6] which is a scripting framework for quantum chemistry implemented in the Python[129] programming language.

6.2.1 BP8/05 data set

For the FDE calculations we used the TW02 non-additive kinetic-energy functional. We employed both a supermolecular (global) expansion basis denoted as FDE(s), in which the basis functions of all subsystems are used for the expansion of the KS orbitals of the active subsystem and a monomolecular basis denoted as FDE(m), in which basis functions located only on the active subsystem are used to expand its KS orbitals. FDE(s) enables one to make a rigorous comparison whereas FDE(m) is used in practice and does not suffer from basis set superposition error (BSSE). Both KS and FDE(s) energies were corrected for BSSE with the counterpoise technique.[201] Unless otherwise noted, we optimised the subsystem

densities from both FDE(m) and FDE(s) calculations in 5 FT cycles. Throughout the text below, FT(i) with $i = 1, 2 \cdots, 5$ denotes the number of employed freeze-and-thaw cycles.

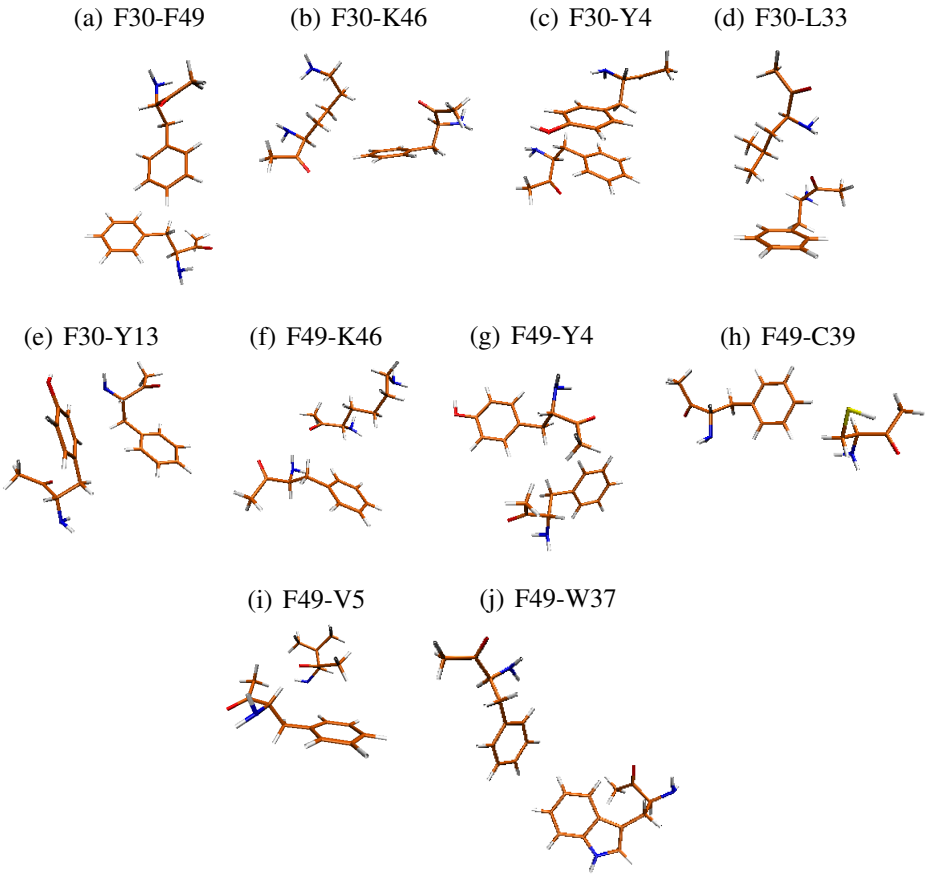


Figure 6.1: Figures of interacting residues in clusters F30 and F49

6.2.2 Rubredoxin

The geometry of Rd is determined by X-Ray crystallography (PDB code 1RB9) with high resolution (0.95 Å), and therefore, geometry optimisation was not necessary. We adopted the approximation employed earlier by Vondrášek *et al.*[192,193] where they excised the hydrophobic-core amino acid residues from the experimental structure and partitioned the whole cluster into two distinct clusters, which are named after the central residues F30 and F49. They further fragmented each cluster into chemically distinct neutral pairs of amino acids that are modelled as

methylated amino acid residues.[193] The central F30 and F49 residues interact with five (F49, K46, L33, Y13, and Y4) and seven (C39, C6, F30, K46, V5, W37, and Y4) amino acids, respectively (see Figure 6.1). We analyzed the interaction energies for this selected set of interacting pairs using the xyz coordinates of the interacting pairs as given in the work of Vondrášek *et al.*[192] We employed the BLYP XC functional and the TW02 and PW86K kinetic-energy functionals. We report here only the results of FDE(m) calculations since we had SCF convergence issues for most of the amino acid pairs with FDE(s) calculations. Similar convergence problems with the FDE(s) scheme were encountered in our earlier work[SMB2] where we showed that the availability of the full basis set for FDE(s), allows the electron density to probe regions where errors in the kinetic energy functional make the embedding potential too attractive.[SMB4] As a consequence, the electron density redistributes to yield strongly overlapping subsystem densities that an approximate functional is not able to describe.

6.3 Results and Discussion

6.3.1 Convergence of Bond Energy Contributions

BP8/05 Data Set In Tables 6.1 and 6.2, we show the convergence behaviour of the bond energy (split up according to Eqs. 6.13 and 6.14 as described in Section 6.1) for the Watson-Crick (WC) hydrogen-bonded base pairs and for the A$\cdots$T dimer as a representative of the stacked base pairs; respectively.

It can be seen that the largest changes in all energy contributions occur during the first three cycles of the FT process. After FT(3), changes are small and at the end of FT(5) both FDE(m) and FDE(s) calculations are fully converged.

As expected for this type of partitioning of the density, the nonadditive kinetic energy term, T_s^{nad}, is positive for both the stacked and hydrogen-bonded base pairs. This unfavourable interaction, that can be interpreted as Pauli repulsion, is opposed by strongly negative potential energy contributions resulting in a negative total interaction energy. Looking at the promotion energy of the fragments we see that the overall positive promotion energy results from a negative kinetic component and positive potential energy term. This is precisely the opposite of what is found in the Morokuma analysis and is a consequence of the fact that orbitals of different subsystems are non-orthogonal and sum to the total density. Whereas the Morokuma scheme proceeds from the orthogonalised fragments to the overall

density via an orbital interaction step, this density deformation towards the bond region is already accounted for in the FDE fragment densities.

For the more strongly bound WC base pairs (Table 6.1) the deviation from the isolated fragment densities that results from the FT-procedure is larger than for stacked base pairs (Table 6.2) but in both cases the bond energies remain rather close to the starting values. The fact that the electrostatic component of the interaction energies become more negative in the course of FT cycles, can be interpreted as mutual polarization of the fragments by the FT procedure. This polarization makes the interaction term more negative, but the effect on the bond energy is minor as this is compensated by an almost equally large increase of the promotion energy.

Rubredoxin In Table 6.3, we show the convergence behaviour of the bond energy (split up according to Eqs. 6.13 and 6.14 as described in Section 6.1) for the amino acid pairs F49$\cdots$V5 as a representative of the cluster around F49 and for F30$\cdots$Y4 representing the cluster around F30. These residues are solely bound by dispersion interactions and DFT, without dispersion correction, yields a positive (repulsive) bond energy that should be reproduced by the FDE approximation.

Similar to DNA base pairs, the energies quickly converge in the FT cycles. In this case, the deviation relative to the starting situation is so small that the bond energy obtained in the first iteration is already sufficiently precise. Interesting is the analysis for F49$\cdots$V5, in which the $\Delta E_{\mathrm{Hartree}}$ and ΔE_{xc} contributions of the promotion energy of F49 are negative whereas the ΔT_s contribution is positive. For F30$\cdots$Y4 and for the base pairs as shown in sub-section 6.3.1 (see Tables 6.1 and 6.2) the signs of these terms are precisely the opposite. This is possibly due to the unusual bonding mechanism in this amino acid pair containing a $\mathrm{CH}\cdots\pi$ contact, which is not present in other pairs tested in this work (for more details on this bonding mechanism see Ref. 192).

Similar to base pairs, for F49$\cdots$V5 the largest changes, in absolute value, occur in the electrostatic contribution to the FDE interaction energy (0.1 kcal/mol) but the magnitude of this changes is much smaller than the 1 kcal/mol found in the base pairs. This indicates that polarization effects are not very important, in agreement with the rapid convergence of the bond energy in the FT cycles.

Also for F30$\cdots$Y4, a structure similar to a parallel-displaced benzene dimer, the electrostatic contribution changes very little during FT cycles. The largest changes occur in the $\Delta E_{\mathrm{Hartree}}$ and ΔT_s components of the promotion energy of F30 but due to their opposite sign these changes do hardly affect the promotion

Table 6.1: BP8/05 data set– WC G$\cdots$C and WC A$\cdots$T: Convergence behaviour of bond energy contributions with FDE(m). FT(i) with $i = 1, 2 \cdots, 5$ denotes the number of freeze-and-thaw cycles with promotion energy and its components for WC G and WC A, promotion energy for WC C and WC T (see Eq. 6.14). WC G$\cdots$C and WC A$\cdots$T denote the interaction energy between two subsystems (see Eq. 6.4) and ΔE_{bond}, the bond energy (see Eq. 6.13). KS is BSSE corrected. All in kcal/mol.

	WC G				WC C	WC G$\cdots$C			
	Hartree	ΔE_{xc}	ΔT_s	E_{prom}	E_{prom}	Electrostatic	$E_{\mathrm{xc}}^{\mathrm{nad}}$	T_s^{nad}	ΔE_{bond}
FT(1)	146.08865	31.50457	−170.73481	6.85841	5.18996	−69.02429	−19.19398	48.39137	−27.77853
FT(2)	146.95411	31.53334	−171.29303	7.19441	6.34303	−70.53423	−19.16633	48.32266	−27.84045
FT(3)	147.02441	31.53671	−171.33715	7.22398	6.43797	−70.65563	−19.16572	48.32125	−27.83815
FT(4)	147.02850	31.53675	−171.34091	7.22434	6.44290	−70.66530	−19.16573	48.32128	−27.84252
FT(5)	147.02879	31.53673	−171.34139	7.22412	6.44472	−70.66611	−19.16574	48.32130	−27.84171
KS									−28.60802
	WC A				WC T	WC A$\cdots$T			
	Hartree	ΔE_{xc}	ΔT_s	E_{prom}	E_{prom}	Electrostatic	$E_{\mathrm{xc}}^{\mathrm{nad}}$	T_s^{nad}	ΔE_{bond}
FT(1)	168.83449	44.76420	−211.49174	2.10695	0.61549	−40.79792	−14.38113	36.45171	−16.00490
FT(2)	169.22939	44.79908	−211.79546	2.23301	1.06093	−41.47270	−14.43176	36.58774	−16.02279
FT(3)	169.25853	44.80220	−211.81754	2.24320	1.09356	−41.52239	−14.43652	36.60040	−16.02175
FT(4)	169.25942	44.80224	−211.82062	2.24105	1.09626	−41.52601	−14.43691	36.60145	−16.02417
FT(5)	169.26023	44.80235	−211.81923	2.24335	1.09574	−41.52632	−14.43694	36.60152	−16.02264
KS									−16.51096

Table 6.2: BP8/05 data set– A$\cdots$T: Convergence behaviour of bond energy contributions with FDE(m) and FDE(s). FT(i) with $i = 1, 2 \cdots, 5$ denotes the number of freeze-and-thaw cycles with promotion energy and its components for A and promotion energy for T (see Eq. 6.14). A$\cdots$T denotes the interaction energy between two subsystems (see Eq. 6.4) and ΔE_{bond}, the bond energy (see Eq. 6.13). FDE(s) and KS is BSSE corrected. All in kcal/mol.

	A				T	A$\cdots$T			
	Hartree	ΔE_{xc}	ΔT_s	E_{prom}	E_{prom}	Electrostatic	$E_{\mathrm{xc}}^{\mathrm{nad}}$	T_s^{nad}	ΔE_{bond}
FDE(m)									
FT(1)	6.06296	2.57751	−8.04578	0.59469	0.75944	−11.16761	−7.22864	19.58776	2.54564
FT(2)	6.01654	2.56304	−7.97976	0.59981	0.79362	−11.18293	−7.22253	19.57468	2.56266
FT(3)	6.01391	2.56230	−7.97678	0.59942	0.79443	−-11.18356	−7.22225	19.57409	2.56214
FT(4)	6.01422	2.56232	−7.97696	0.59959	0.79241	−11.18359	−7.22224	19.57407	2.56024
FT(5)	6.01419	2.56232	−7.97697	0.59954	0.79550	−11.18360	−7.22224	19.57406	2.56327
FDE(s)									
FT(1)	9.44623	3.01111	−12.08908	0.36826	0.55798	−12.21583	−7.15233	19.74805	2.14936
FT(2)	9.40639	2.99927	−12.03257	0.37310	0.59023	−12.23865	−7.14999	19.74227	2.16020
FT(3)	9.40386	2.99860	−12.02978	0.37268	0.59117	−12.23957	−7.14982	19.74190	2.15959
FT(4)	9.40422	2.99863	−12.02998	0.37287	0.58874	−12.23962	−7.14982	19.74188	2.15729
FT(5)	9.40418	2.99863	−12.03000	0.37280	0.59238	−12.23963	−7.14982	19.74188	2.16085
KS									3.59592

energy.

6.3.2 Accuracy of Bond Energies with Dispersion Correction

BP8/05 Data Set Table 6.4 shows the results for bond energies obtained with KS-D, FDE(s)-D, and FDE(m)-D together with the dispersion correction contribution D as well as reference values[154,202,203] obtained with accurate wave function based methods. As the KS-D results show, the dispersion correction gives a significant improvement of the interaction energies for the stacked base pairs, with the exception of par C$\cdots$C. Because the KS contribution is well reproduced by FDE, basically the same results are obtained with FDE-D at a lower cost.

The BLYP XC functional was found by Grimme[162] to be the most suitable when applied with a dispersion correction. For hydrogen bonded complexes one may, however, experience some "double-counting" of correlation effects, since in these complexes DFT can already partly account for dispersion contributions due to the overlapping fragment electron densities. This explains the overestimation of the interaction energies seen in both KS-D and FDE(m)-D in the last two lines of Table 6.4. This can be remedied by a better balancing of the XC and dispersion energy terms in KS-D and FDE-D and does not affect our comparison of these two methods.

Rubredoxin Table 6.5 collects the KS-D, FDE(m)-D results and dispersion correction D as well as reference interaction energies obtained with resolution of identity second-order Mϕller Plesset method extrapolated to the basis set limit.[192] In this case inclusion of dispersion interactions is mandatory to obtain a qualitatively correct result. With dispersion, the KS results are in a reasonable agreement with the reference energies with the correct sign and with deviations of the order of a few kcal/mol. Given this accuracy, the differences between KS-D and FDE-D are acceptable with a largest discrepancy of 1.3 kcal/mol for the most strongly interacting pair (F49 – V5). As noted earlier, one should not compare FDE–D directly with the reference MP2 results because this would lead to the conclusion that FDE(m)-D outperforms KS-D in some cases. This is due to error cancellation because FDE(m)-D is a well-defined approximation of KS-D which, in the limit of an exact T_s functional, should reproduce the KS-D result.

Table 6.3: Rubredoxin– F49$\cdots$V5 and F30$\cdots$Y4: Convergence behaviour of bond energy contributions with FDE(m). FT(i) with $i = 1, 2 \cdots, 5$ denotes the number of freeze-and-thaw cycles with promotion energy and its components for F30 and F49, promotion energy for Y4 and V5 (see Eq. 6.14). F30$\cdots$Y4 and F49$\cdots$V5 denote the interaction energy between two subsystems (see Eq. 6.4) and ΔE_{bond}, the bond energy (see Eq. 6.13). KS is BSSE corrected. All in kcal/mol.

	F49				V5	F49$\cdots$V5			
	Hartree	ΔE_{xc}	ΔT_s	E_{prom}	E_{prom}	Electrostatic	$E_{\text{xc}}^{\text{nad}}$	T_s^{nad}	ΔE_{bond}
FT(1)	−4.51665	−3.06945	9.07307	1.48698	0.43383	−11.01564	−7.40670	18.53846	2.03693
FT(2)	−4.50650	−3.06902	9.07665	1.50114	0.47983	−11.07774	−7.40889	18.54293	2.03728
FT(3)	−4.50536	−3.06887	9.07628	1.50205	0.48203	−11.08104	−7.40903	18.54322	2.03725
FT(4)	−4.50494	−3.06880	9.07677	1.50304	0.48368	−11.08122	−7.40903	18.54324	2.03971
FT(5)	−4.50469	−3.06875	9.07601	1.50257	0.48338	−11.08123	−7.40903	18.54324	2.03893
KS									3.31354
	F30				Y4	F30$\cdots$Y4			
	Hartree	ΔE_{xc}	ΔT_s	E_{prom}	E_{prom}	Electrostatic	$E_{\text{xc}}^{\text{nad}}$	T_s^{nad}	ΔE_{bond}
FT(1)	2.54430	0.46561	−2.56001	0.44990	0.36399	−3.29746	−2.86293	8.53177	3.18526
FT(2)	2.51306	0.45618	−2.51943	0.44981	0.38126	−3.29988	−2.85939	8.52475	3.19656
FT(3)	2.51163	0.45581	−2.51747	0.44997	0.38122	−3.29998	−2.85926	8.52449	3.19644
FT(4)	2.50997	0.45549	−2.51683	0.44863	0.38183	−3.29998	−2.85925	8.52448	3.19571
KS									3.76713

Table 6.4: BP8/05 data set: KS–D, FDE(s)–D (BSSE-corrected), FDE(m)–D bond energies and dispersion correction D with the BLYP XC functional in kcal/-mol. FDE bond energies are with TW02 kinetic energy functional. Ref. refers to reference interactions energies obtained with wave-function based methods. For hydrogen bonded complexes and anti C$\cdots$C, no SCF convergence could be reached for FDE(s).

Complexes	D	KS-D	FDE(s)-D	FDE(m)-D	MP2
Stacked					
A$\cdots$T	−15.02	−11.42	−12.85	−12.45	−11.60 [202]
G$\cdots$C	−11.98	−15.19	−16.92	−16.59	−16.90 [202]
U$\cdots$U	−10.27	−8.86	−9.73	−9.14	−10.30 [154]
par C$\cdots$C	−8.99	−3.06	−5.26	−5.03	2.45 [203]
anti C$\cdots$C	−10.32	−10.25	—	−10.99	−9.90 [154]
displ C$\cdots$C	−8.64	−9.23	−9.98	−9.79	−9.43 [203]
Hydrogen bonded					
WC A$\cdots$T	−5.28	−21.80	—	−21.31	−15.40 [204]
WC G$\cdots$C	−6.31	−34.92	—	−34.15	−28.80 [204]

6.3.3 Efficiency of FDE(m)-D

In order to assess the efficiency of FDE(m)-D, we measured the duration of each calculation and compared it with that of KS-D. For the F30 cluster, the KS-D calculations took 48 to 74 minutes, whereas the 5 FT cycles of FDE-D took 33 to 52 minutes. For the full calculation, FDE(m)-D was about 30% more efficient than KS-D. This difference was slightly smaller for the F49 cluster for which FDE(m)-D was only 24% more efficient than KS-D. We did not time a case with only 3 FT-cycles, but given the linear dependence of the calculation time on the number of FT cycles, one may easily estimate that such FDE calculations would be about a factor of 3 faster than KS calculations. This efficiency can be further increased by improving upon the costly evaluation of the fit corrections to the Coulomb interaction energy. The current FDE implementation in ADF is primarily optimised for molecular property calculations in which evaluation of this term is not needed.

Table 6.5: Rubredoxin– pairs clustered around F30 and F49: KS–D (BSSE corrected), FDE(m)–D, dispersion correction D with the BLYP XC functional in kcal/mol. Ref. refers to reference interaction energies obtained with resolution of identity second-order Mϕller Plesset method using a complete atomic orbital basis set limit.[192]

	D	KS-D	FDE–D	MP2
F30				
F49	−4.15	−2.47	−2.30	−3.30
K46	−3.49	−2.86	−2.58	−3.40
Y4	−8.97	−5.21	−5.78	−7.00
L33	−8.15	−5.32	−5.76	−5.50
Y13	−5.10	−2.28	−2.72	−4.50
F49				
K46	−5.83	−4.49	−4.36	−4.80
Y4	−7.43	−3.31	−2.80	−3.10
C39	−4.37	−2.34	−2.11	−2.10
V5	−9.48	−6.17	−7.44	−6.70
W37	−3.01	−1.74	−1.79	−2.50
C6	−7.69	−3.72	−2.66	−5.00

6.4 Conclusions

We propose density functional theory including dispersion correction as approximated by the frozen-density embedding method (FDE-D) as an efficient method to calculate interaction energies between biomolecular fragments. Results for selected test cases using the dispersion correction of Grimme[167] in conjunction with the BLYP XC functional and the TW02 kinetic-energy functional indicate that sufficient accuracy can be reached at a significantly lower cost compared to supermolecular DFT calculations.

In order to analyse the bonding between the fragments in chemically meaningful terms, we propose a bond energy decomposition inspired by the KS bond formation analysis[28] but allowing for nonorthogonality of orbitals that belong to different subsystems. We can thereby write the bond energy in terms of promotion energies of individual subsystems and an explicit interaction energy between these (Eq. 6.13). This analysis provides a new view on the effect of Pauli repulsion and has the unique feature of being done in terms of interacting fragment densities that sum to the full molecular density.

For all the molecules tested in this work, we show that the electrostatic attraction between the subsystem electron densities increases in the course of the FT process which confirms the intuitive picture of FT providing mutual polarization.

Chapter 7

Computational study on the anomalous fluorescence behavior of isoflavones [1]

Abstract

Isoflavones are known to show fluorescence with intensities that depend strongly on the solvent properties and exhibit Stokes' shifts as large as 1.4 eV. While some of this behaviour can be explained by (excited state) deprotonation, this mechanism does not apply for all isoflavones. In this chapter, a computational and experimental investigation of the reasons for this anomalous behaviour of neutral isoflavones, taking the daidzein molecule as a model compound is given. It is shown that the absence in fluorescence in aprotic solvents and the weak fluorescence in protic solvents can be explained by a change of order of the lowest singlet states in which a fluorescent charge-transfer state lies below the nonfluorescent locally excited state in water but not in acetonitrile. The large Stokes' shift is partly due to a significant rotation among the chromone-phenyl bond in the excited state.

[1] The contents of this chapter have been published as S. M. Beyhan, A. W. Götz, F. Ariese, L. Visscher, and C. Gooijer *J. Phys. Chem.*,**115**, 1493 (2010)

7.1 Introduction

Due to anti-oxidant activities[205–207] and associated beneficial effects on human health[208–211] flavonoids have received considerable attention and their properties have been studied by various experimental techniques. An interesting observation is the fact that the fluorescence properties of flavonoids in solution are often strongly dependent on the nature of the solvent. This property makes them also interesting as probes in biochemical studies,[212] and has led to considerable interest in the electronically excited states of flavonoids. Much attention has been given to flavones, the most widely occurring class of flavonoids. Well-known among these are the 5-hydroxy flavones (5HF), which were initially thought to be non-fluorescent,[213,214] but with the help of laser excitation were shown to have in fact dual emission bands.[215] This could be explained by excited-state intramolecular proton transfer (ESIPT),[216] leading to large solvent effects. Also the 3-hydroxyflavones (3HF) and the 7-hydroxyflavones (7HF) exhibit large Stokes' shifts attributed to ESIPT.[217–226] For 7HF, in addition to ESIPT, the stabilization of the S_1 excited state by Coulombic interactions with polar solvents plays an important role.[224,225,227] In a recent computational study concerning the unsubstituted flavone molecule, Marian[228] showed that the first excited singlet state favors a planar arrangement of the phenyl and chromone rings, in contrast to the twisted conformation adopted in the electronic ground state. In the same study it was shown that the intersystem crossing rate (ISC) for the $S_1' \rightsquigarrow T_1$ transition, where S_1' denotes the relaxed S_1 state, is significant while decreasing with solvent polarity and protic character.

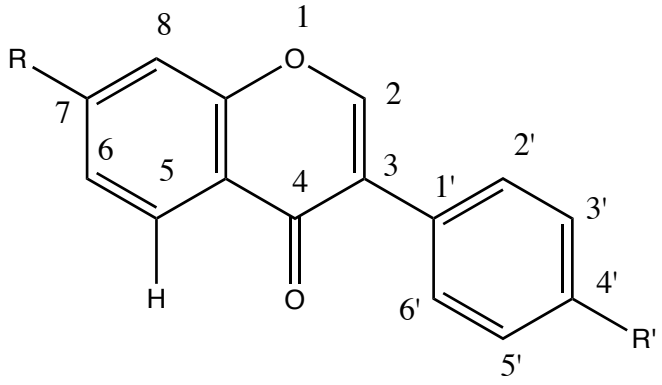


Figure 7.1: General structure of isoflavones (3-phenylchromen-4-ones). For daidzein R=OH, R′=OH

A less common subclass of the flavonoids, which is naturally found in soy

beans and other plants of the Leguminosae family,[229] is formed by the isoflavones (3-phenylchromen-4-ones) (Figure 7.1). Wolfbeis and coworkers[230,231] and De Rijke and coworkers[232] studied a range of isoflavones and reported that some of the natively fluorescent compounds exhibit a very large Stokes' shift ($\approx$ 11 000 cm^{-1} or 1.4 eV, determined in this paper as the difference between the maxima of the longest-wavelength absorption band and the fluorescence emission) that is very sensitive to the polarity of the solvent. For these molecules the large Stokes' shift cannot be attributed to ESIPT because there are no acidic hydrogens in the immediate vicinity of the basic carbonyl group at C4. For isoflavones with an -OH substituent at C7 strong fluorescence was observed under basic conditions, which was attributed to the anionic form.[230,232,233] It was suggested that in the case of 7-hydroxyisoflavone (R = OH; R' = H in Figure 7.1) the emission at neutral pH can also occur from the excited-state anion due to the increased acidity of the 7-OH group in the excited state. On the other hand, the excitation spectra presented by Dunford *et.al* on daidzein (R = OH; R' = OH)[233] at neutral pH indicated that the emission was mainly due to a small population of deprotonated molecules in the ground state. Surprisingly, fluorescence with a large Stokes' shift was also observed in the case of ononin (R = O-glucose; R' = $\mathrm{OCH_3}$ in Figure 7.1), a molecule that cannot be deprotonated at C7 and shows practically no pH dependence.[232] Furthermore, in the case of daidzein fluorescence was also observed under acidic conditions. These findings show that deprotonation is not required and that fluorescence can also occur in neutral isoflavones. The large Stokes' shift could then originate from a large difference in polarity between ground and excited state.[233] Interestingly, there is little or no fluorescence in polar but aprotic solvents such as acetonitrile (ACN).[232,233] In the case of ononin an unusual decrease in fluorescence was observed at low temperatures, and it was postulated that a major geometrical change is required for emission to occur.[232] To explain these findings it is natural to employ computational techniques to verify whether the premises of this interpretation, large differences between the ground and excited state geometries and charge distributions, are indeed correct.

We chose to study the simplest isoflavone, the neutral daidzein molecule, as a representative example that exhibits the essential features of this class of molecules. In order to discuss the effect of water as a solvent we measured the fluorescence spectrum in water to check for differences with the spectrum recorded earlier by Dunford[233] in pure methanol. The computational study on this compound has two objectives. First, to establish the amount of structural relaxation to be expected after excitation and, second, to determine how solvents influence the energy gaps between the ground state and the singlet manifold of states, as well as the position of the lowest triplet states and thereby the likelihood of intersystem crossing. In addition to excited state studies in the gas phase, we used a

microsolvation model to simulate the effects of both a polar protic solvent and a polar aprotic solvent, using respectively one, three or four water molecules or four acetonitrile molecules to describe the most important solute-solvent interactions.

The most commonly employed method to study excited states of molecules of this size is time-dependent density functional theory (TDDFT).[234–236] Since it is difficult to assess the reliability of TDDFT excitation energies when charge-transfer effects can play a role we have chosen to use the second-order approximate coupled cluster approach (CC2)[237] for the majority of our calculations. We had to rely on TDDFT only for excited state geometry optimizations of our computationally most extensive microsolvation models which turned out to be too complex for CC2. Similar studies on coumarin dyes have been reported in the literature showing that TDDFT with hybrid functionals can yield reasonable agreement with experimental excitation energies and structures.[238–243] We chose to use the hybrid functional of Perdew, Burke, and Erzerhof (PBE0)[244,245] for the DFT calculations reported here and have assessed the reliability of these calculations by comparison to our CC2 results for the gas phase model in which both types of calculations are feasible.

In Section 7.2, we will present details of the experimental and computational methods that we have employed. Section 7.3 will be devoted to a discussion of the results for excited states, organized by subsections treating first the experimental data, then the calculations on the isolated daidzein molecule and subsequently the microsolvation models. Triplet states are briefly discussed in the final Section 7.3.5, and we conclude with a summary and short outlook.

7.2 Computational and experimental details

All computations were carried out with version 5.10 of the TURBOMOLE program package.[246,247] The PBE0 exchange-correlation (XC) functional,[244,245] standard TURBOMOLE numerical XC quadrature grids,[248] and a triple-ζ valence-polarized contracted Gaussian basis set (TZVP)[249] were used for DFT calculations of ground state properties and equilibrium geometries and for TDDFT calculations of singlet and triplet excitation energies[250] as well as excited state properties and equilibrium geometries.[251,252]

The aug-cc-pVDZ basis set[253,254] with corresponding auxiliary basis set[255] was used for parallel[256] CC2[237] calculations of ground state properties and equilibrium geometries,[257] vertical singlet and triplet excitation energies,[258,259] transition mo-

ments and excited state first-order properties,[260,261] as well as excited state equilibrium geometries[262] and interaction energies between the acetonitrile molecules and daidzein. All 1s orbitals were frozen and thus excluded from the correlation treatment and excitation process.

Accurately converged self-consistent field (SCF) results (a termination threshold of 10^{-8} au for the total energy and 10^{-7} au for the density matrix) were used to guarantee satisfactory accuracy in the TDDFT calculations and CC2 calculations. Geometry optimizations were considered converged if the maximum component of the Cartesian gradient dropped below 10^{-4} au.

Daidzein (ICN Biomedicals) solutions of 1 x 10^{-5} M were prepared in acetonitrile (ACN), and in water/methanol mixtures 90:10 v/v of which the water phase had been adjusted to the listed pH (daidzein is not sufficiently soluble in water at low pH). UV-Vis absorption spectra were recorded in 1-cm pathlength quartz cuvettes, on a Shimadzu UV2501 dual beam spectrometer at 2 nm spectral resolution. Uncorrected fluorescence emission and excitation spectra were recorded using a Perkin Elmer LS-50B spectrometer and excitation/emission spectral slit widths of 10/10 nm. A 390 nm emission cut-off filter was fitted to reject second-order scattering interferences.

7.3 Results and discussion

7.3.1 Absorption and fluorescence spectra

Figure 7.2a shows the absorption spectrum of neutral daidzein in acetonitrile and in water/methanol 90:10 (pH = 3). In pure water at pH = 3 the compound is poorly soluble, and the absorbance was weak but otherwise identical (not shown). The spectrum is very similar to the absorption spectrum presented by Dunford *et.al* for methanol solutions.[233] The longest-wavelength absorption band of daidzein occurs at about 305 nm, whereas the strongest absorbance occurs at 248 nm in water and at 260 nm in acetonitrile. The same bands at 248 nm and 305 nm are observed in the fluorescence excitations spectrum depicted in Figure 7.2b; the emission maximum is at 460 nm. Daidzein is only weakly fluorescent in aqueous solutions under acidic conditions. No significant emission was observed in acetonitrile. In all solvent systems tested, the emission is strongly red-shifted relative to the absorption band at 305 nm, and also strongly red-shifted relative to the 337 nm lowest absorption band of the anion.[233] The new measurements in

aqueous solution are in line with the red shifting of the emission maximum with increasing solvent polarity.[233]

7.3.2 Absorption and emission without solvent effects

By means of calculations we will now start to characterize the excited state energies, fluorescence intensities and solute-solvent interactions in order to explain the large (1.4 eV) Stokes shift and the variation of the fluorescence intensities with the solvent. We will focus on the neutral (not deprotonated) daidzein molecule and first look at calculations in which solvent effects are not included.

For the ground state geometry of the isolated daidzein molecule, the CC2 method predicts a dihedral angle (ϕ=C2-C3-C1′-C2′, compare Figure 7.1) of $39.9°$ between the chromone and the phenyl ring. Vertical excitation energies, oscillator strengths and dipole moments at this ground state geometry are given in Table 7.1. The fifth excited state (S_5) has by far the highest oscillator strength and its CC2 excitation energy of 5.09 eV matches very well with the spectra in Figure 2a that show an absorption maximum of 5.0 eV (248 nm) for daidzein in water and 4.9 eV in ACN.

The states relevant for the emission are the two lowest excited singlet states, of which the first, S_1, mainly results from a local excitation (LE) on the chromone rings, whereas the second excited singlet state S_2 can be characterized as a charge transfer (CT) state in which charge is moved from the phenyl to the chromone. As the order of the states changes in the different computational methods, in the following discussion these states will be referred to as LE and CT; respectively. The calculated energy of the LE state (3.95 eV = 313 nm) corresponds well with the absorption spectra of Figure 7.2 (305 nm). In the water models, the LE excitation energy varies between 4.10 and 4.12 eV, which also fits well with experimentally observed maximum of 4.10 eV in water. A problem with this assignment is, however, the negligible oscillator strength of the LE state in both the gas phase and the microsolvated models. Given the fact that the error in the CC2 method can amount to 0.3 eV,[263] it is also possible that the higher-lying CT state with the more significant oscillator strength of 0.098 is responsible for the lowest absorption band.

In order to study the experimentally observed emission maximum of about 460 nm (2.7 eV) in more detail we need to consider structurally relaxed states. The corresponding adiabatic excitation energies and excited state dipole moments are shown in Table 7.2, together with a characterization of the excited state structures in terms of the dihedral angle ϕ between the rings.

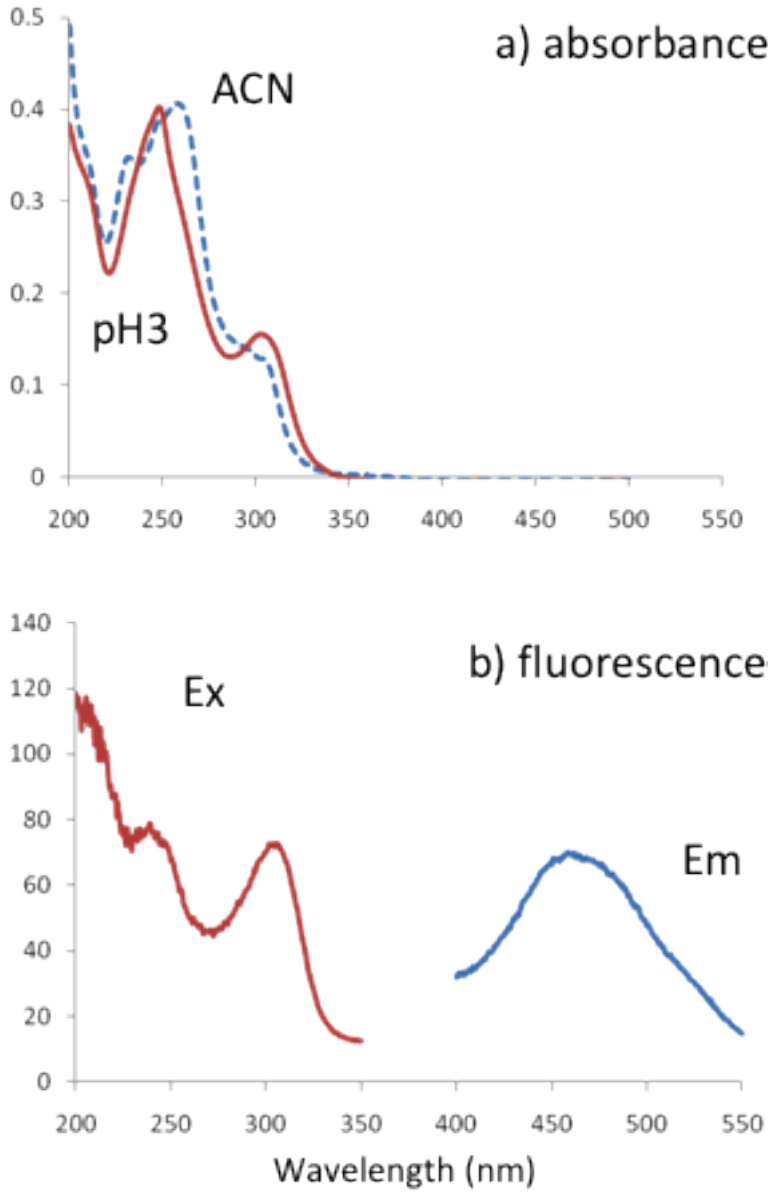


Figure 7.2: a) Absorption spectra of 10^{-5} M daidzein in acetonitrile (dashed line) and in water/methanol 90:10 v/v at pH = 3 (full line); optical pathlength = 1.00 cm. b) Uncorrected fluorescence excitation (λ_{ems} = 465 nm) and emission spectra (λ_{exc} = 303 nm) of 10^{-5} M daidzein in water/methanol 90:10 v/v at pH = 3; intensities in arbitrary units.

Table 7.1: Vertical excitation energies ΔE_{abs} (eV), oscillator strengths, and dipole moment μ (Debye) of isolated and microsolvated daidzein (calculated with the CC2 method).

State	no solvent			$1H_2O$		$3H_2O$		$4H_2O$		4ACN	
	ΔE_{abs}	f	μ	ΔE_{abs}	f	ΔE_{abs}	f	ΔE_{abs}	f	ΔE_{abs}	f
S_0			2.47								
LE	3.95	0.000	2.49	4.10	0.000	4.12	0.000	4.10	0.000	4.08	0.000
CT	4.43	0.098	4.58	4.38	0.081	4.31	0.068	4.28	0.065	4.41	0.095
S_3	4.68	0.002	1.39	4.63	0.020	4.55	0.056	4.54	0.073	4.57	0.072
S_4	4.78	0.030	2.38	4.79	0.032	4.73	0.033	4.74	0.035	4.69	0.020
S_5	5.09	0.544	2.74	5.08	0.726	4.92	0.226	4.93	0.568	5.02	0.758

Table 7.2: Dihedral angles ϕ (degrees, ϕ =C2-C3-C1′-C2′, see Figure 7.1), emission energies ΔE_{ems} (eV), oscillator strengths f and dipole moments μ (Debye) of the CC2 optimized ground and excited states of isolated daidzein.

State	ϕ	ΔE_{ems}	f	μ
S_0	39.9			2.47
LE	33.3	1.99	0.001	3.26
CT	12.6	3.40	0.122	5.70

Table 7.2 shows an appreciable rotation towards planarity along the chromone-phenyl bond for the relaxed CT state in which the dihedral angle is reduced to $12.6°$. This can be explained by a reduction of π^*-character in the chromone-phenyl bond which increases the bond order and drives the chromone and phenyl rings towards a planar arrangement. The dipole moment in this relaxed CT structure is 5.70 D, which is more than twice the ground state value of 2.47 D. This is in line with the hypothesis of Dunford[233] who proposed emission from a state that is stabilized by polar solvent. Also the significant oscillator strength points at this state as being responsible for emission. A problem with this explanation is the fact that the difference between this state and the lower LE state is now greatly increased. The relaxed LE state, which still has a very small oscillator strength, is calculated at 1.99 eV above the ground state surface. This state is characterized by a significant weakening and strong elongation of the carbonyl (C4-O, compare Figure 7.1) bond. For both states one may therefore expect a significant influence of solvent molecules.

Since CC2 excited state structure optimizations were computationally not feasible for microsolvation models, we carried out these optimizations using the TD-DFT (PBE0) method. As the TD-DFT calculations may change the ordering of states we need to establish the correspondence between the states calculated in the two computational models. This is readily done by plotting the charge density difference between ground and excited states for both methods. These plots may be found in the supporting information 10 and indicate a reversal of the order of the two lowest states between the two methods, in which the CT state comes out as the lowest excitation when calculated with PBE0 at the ground state geometry. The second excited state calculated with the PBE0 method does resemble the CC2 LE state although there are some differences visible in the two density difference plots for this state. To test the assignment of the CT state we performed a CC2 calculation at the PBE0 S_1 structure and obtained an excitation energy of 3.35 eV that is close to the excitation energy calculated at the fully optimized CC2 structure (3.40 eV). This corroborates the identification of this pair of states and

indicates that the PBE0 structure is sufficiently close to the CC2 optimized structure. For the other pair of states, the calculations are more problematic as it was difficult to fully converge the PBE0 structure optimization for the second excited state. With a milder treshold for convergence of 10^{-3} au, we obtained a structure with a less extended carbonyl bond and a much higher excitation energy for the LE state then when computed at the CC2-optimized structure.

We will return to this identification of states when discussing the microsolvation results.

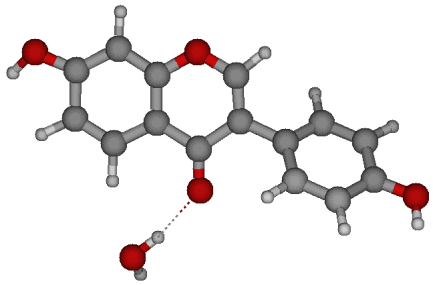
7.3.3 Effect of microsolvation on the absorption

In order to model the effects of protic and aprotic solvents, we next consider microsolvation models with different numbers of water or acetonitrile molecules. CC2 optimized ground state structures of these microsolvated systems are shown in Figures 7.3 and 7.4.

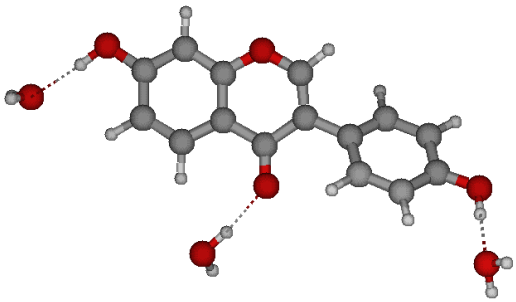
Microsolvation with water The initial position of the water molecules was chosen such that only the waters with a specific hydrogen bonding interaction to the solute are included. This makes it possible to detect the effect of a specific hydrogen bond. For example, in the smallest microsolvation model we only include the water bonding to the carbonyl oxygen as this is likely to have the largest effect on the excited state properties.

As illustrated in Figure 7.3, microsolvation does not strongly affect the optimal twisting angle between the rings for the ground state structure, yielding $41.7°$, $39.7°$ and $39.4°$ for one, three and four water molecules, respectively. This is very similar to the $39.9°$ of the isolated molecule (see Table 7.2). The vertical excitation energies show small solvatochromic shifts that range from -0.16 to 0.15 eV, with the LE state being shifted to the blue (Table 7.1) whereas the CT state is shifted to the red. The oscillator strengths change somewhat upon microsolvation but the difference between S_5 and the other states remains large. As can be seen from Table 7.1, the largest changes in the excitation energies and the oscillator strengths occur upon going from the isolated daidzein molecule to a microsolvation model with one and three water molecules, respectively. Addition of the fourth water molecule does not result in a significant change, indicating that three water molecules – if positioned correctly – are sufficient to capture the most significant interactions at the ground state geometry.

(a) $1H_2O$



(b) $3H_2O$



(c) $4H_2O$

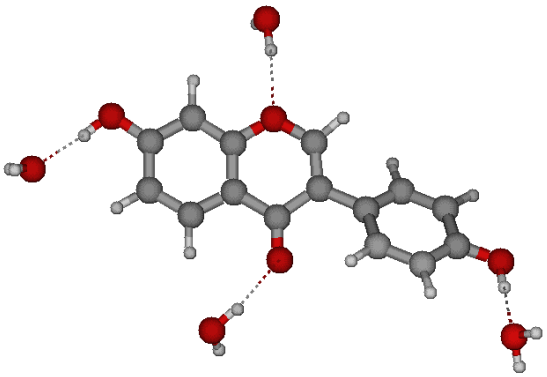


Figure 7.3: Microsolvated daidzein with different numbers of water molecules

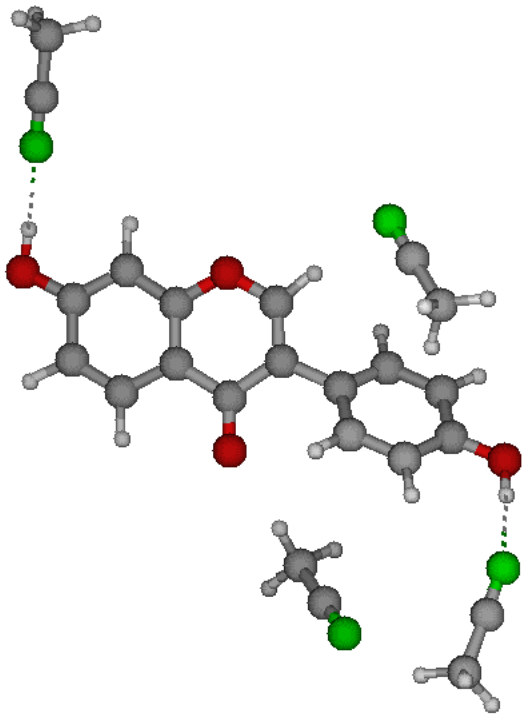


Figure 7.4: Microsolvated daidzein with four acetonitrile molecules

Microsolvation with acetonitrile As a representative of a polar, aprotic solvent we considered a microsolvated system containing four acetonitrile molecules. The CC2 ground state optimized structure of this system is shown in Figure 7.4. The dihedral angle ϕ between the rings is $44.1°$, which is similar to the results in the gas phase and with the water microsolvation models. The corresponding singlet excitation energies and oscillator strengths are also listed in Table 7.1. As expected, solvation by acetonitrile results overall in somewhat smaller solvatochromatic shifts than found in the water microsolvation models, ranging from -0.09 to 0.13 eV with again the LE state being shifted to the blue and the CT state to the red. The S_5 oscillator strength is similar to the water and gas phase results. Only two ACN molecules have significant interactions with the -OH groups (17 kcal/mol), the other two shown in Figure 7.4 have only very little interaction (2 kcal/mol). For the time-consuming excited state optimizations it is thus sufficient to take only two ACN molecules into account.

7.3.4 Effect of microsolvation on the emission

To compute the emission spectra, we perform CC2 single point calculations at the PBE0-optimized structures. To check the correctness of the identification of the the CC2 and PBE0 states, we again computed density difference plots at the ground state structure for both methods. These plots may be found in Figure 2 of the supporting information and indicate the reversal of the ordering of the two lowest states between the two methods, with the microsolvated CC2 LE state now more clearly resembling the PBE0 S_2 state, while the CC2 CT state is again identified as the PBE0 S_1 state, thus confirming the identification proposed in the gas phase model.

The effect of microsolvation on the length of the carbonyl bond (C4-O, compare Figure 7.1) in the relaxed LE and CT states is displayed in Table 7.3 and we note for the LE state in the D+3H_2O model the lengthening relative to the S_0 value of 123 pm is even more pronounced in the microsolvated models. Of more interest is the distance of the proton of the water coordinating to the carbonyl. For the CT state, this distance is decreased by 0.1 Å relative to the distance found in the ground state, while for the LE state we find an elongation by 0.1 Å.

In the D+3H_2O model both PBE0 and CC2 now give the CT state as the lowest excited state when calculated at its optimized geometry. The LE state still represents a lower emission energy (3.26 eV) at its own optimized geometry but the difference of 0.23 eV is much smaller than calculated in the isolated molecule. This value is in the range of errors that can be expected from the CC2 model and

Table 7.3: Carbonyl bond lengths (pm) (C4-O, compare Figure 7.1) of isolated and microsolvated daidzein at the PBE0 optimized ground state and first two excited state structures.

State	no solvent	$3H_2O$	2ACN
S_0	122	123	123
LE	128	131	129
CT	128	127	126

Table 7.4: CC2 emission energies ΔE_{ems} (eV) and oscillator strengths f at the PBE0 optimized excited state microsolvation structures. D refers to the isolated daidzein molecule, D+$3H_2O$ and D+2ACN refer to microsolvated systems.

State	D		D+$3H_2O$		D+2ACN	
	ΔE_{ems}	f	ΔE_{ems}	f	ΔE_{ems}	f
LE	3.35	0.000	3.26	0.001	3.32	0.000
CT	3.89	0.061	3.49	0.034	3.70	0.042

the neglect of long range polarizations. If we assume emission and absorption from the CT state this would yield a Stokes shift of 4.31-3.49 = 0.82 eV, while taking LE as lowest absorbing state yields the slightly smaller value of 4.12-3.49 = 0.63 eV. For the acetonitrile this situation is different as the absence of a coordinating hydrogen bond gives a weaker solute-solvent interaction which is too small to overcome the difference in excitation energy. In this case the CT state is only lowered to 3.70 eV and remains the second excited state also at its optimized geometry.

For the LE state the influence of the solvating water molecules is smaller, but again pointing to a somewhat stronger stabilization of this state (relative to the isolated molecule) than in the acetonitrile microsolvation model (see Table 7.4). This is consistent with the fact that the dipole moment of the LE state is slightly larger than that of the ground state (Table 7.2). In both cases, the oscillator strength is not much affected, coming for the water model now slightly above the threshold of three digits that we chose for displaying the results.

In conclusion of this part, we find a significant influence of the presence of the solvent molecules on the excitation energies, the water model yielding a Stokes'

Table 7.5: Vertical CC2 singlet and triplet excitation energies (eV) for isolated daidzein (ID) and microsolvation models of daidzein with 4 water molecules (D+4H_2O) and 4 acetonitrile molecules (D+4ACN).

State	ID	D+4H_2O	D+4ACN
LE	3.95	4.10	4.08
T_1	3.56	3.34	3.62
T_2	3.73	3.80	3.88
T_3	4.02	3.91	3.96
T_4	4.10	3.97	4.08
T_5	4.47	4.36	4.36

shift of 0.82 eV. In agreement with the experimental data, this model predicts the fluorescence of the neutral species to be weak in water and absent in acetonitrile. In order to check also other possible explanations for the observed absence of fluorescence in ACN, we will next turn to the calculations of the triplet excited states.

7.3.5 Triplet states

The negligible fluorescence quantum yield in polar, aprotic solvents could also be attributable to effective intersystem crossing (ISC). An important parameter determining its efficiency is the energy difference between the singlet manifold and low-lying triplet states, although other parameters play a role as well. In order to obtain an indication for the efficiency of ISC we calculated the position of the lowest triplet states. As a full analysis of the potential energy surfaces is beyond the scope of the current paper, we will restrict ourselves to a comparison of the first singlet and triplet vertical excitation energies in the gas phase (D) and the microsolvation models with four water molecules (D+4H_2O) and four acetonitrile molecules (D+4ACN) computed at the CC2 level (Table 7.5). From the results at this geometry it is evident that addition of water increases the energy gap between LE and the triplet states considerably. In the gas phase and acetonitrile models the energy gap between the two manifolds is small, making ISC more likely. In acetonitrile, ISC could thus also be a cause of the observed absence of fluorescence.

7.4 Conclusions

The new experimental data agree with the older measurements on related compounds in finding a weak aqueous fluorescence with a large Stokes' shifts of daidzein under acidic conditions and no significant emission in acetonitrile. Calculations of the vertical excitation energies agree well with the absorption spectra and indicate that the strongest absorption band corresponds with the fifth excited state.

Relaxation of the lowest excited singlet states is always accompanied by a significant rotation towards planarity around the chromone-phenyl bond relative to the ground state structure which has a dihedral angle $\phi = 39.9°$. The largest change thereby occurs in the CT state, which has a dihedral angle of only $\phi = 12.6°$ at its equilibrium geometries. In the D+3H_2O PBE0 model we find a crossing of the LE and CT surfaces that is absent in the isolated molecule and acetonitrile models. This supports the observation that interaction with protic solvents induces fluorescence of daidzein also under acidic conditions. The relaxation to the emitting CT state is accompanied by a rotation among the chromone-phenyl bond to an almost planar arrangement of all three rings as was already anticipated by de Rijke when studying the temperature dependence of the fluorescence.[232]

According to the spectra of Figure 2b, the emission occurs at around 460 nm (2.7 eV), which even if we allow for an error due to the neglect of vibrational broadening is a significantly lower energy than the lowest emission found in our best water model (3.49 eV). This may partly be caused by artifacts of the computational protocol (CC2 at TD-DFT structures). We furthermore note that models that allow for more extensive solvent polarization should also increase the Stokes' shift given the larger dipole of the CT state compared to the ground state dipole moment. Already in our small solvent model, the difference in interaction between protic solvents (as exemplified by water) and aprotic solvent (acetonitrile) is, however, sufficient to reverse the order of the LE and CT states. The CT is thereby appearing as a (weakly) fluorescent state in water, while the daidzein in acetonitrile decays via the nonfluorescent LE state. Qualitatively this model can therefore explain all experimental observations. We note that calculating the emission energies and the Stokes' shift more accurately requires excited state optimizations at a higher level of theory and the use of a larger solvation model. This is planned for future work in which we intend to also treat related molecules.

Finally, we have shown that the energy gap between the singlet manifold and the lowest triplet states is smaller in acetonitrile compared to water. It may be of interest to study ISC rates in more detail as this may provide an alternative

explanation for the complete lack of fluorescence observed in aprotic solvents.

Chapter 8

Summary

Subsystem density-functional theory (DFT) is an efficient alternative to the conventional Kohn-Sham (KS) method as it allows an accurate description of the interaction between the subsystem of interest and the environment with shorter computer time. This interaction is represented by an effective embedding potential, that contains the electrostatic potential of the environment, an exchange-correlation component and a kinetic-energy component. The formulation of subsystem DFT is in principle exact unlike most other embedding schemes in theoretical chemistry. Owing to these advantages, subsystem DFT is a scheme that shows indication of future excellence to study large systems where the balance of efficiency and accuracy is of great importance.

This thesis investigates the present shortcomings, and their reasons, of subsystem DFT by a thorough analysis of the effective embedding potential. It gives a particular focus to the behavior of the effective embedding potential in the presence of which the electron density of the subsystem of interest is determined. Moreover, it further evaluates the applicability of subsystem DFT.

Chapter 2 gives a review of DFT with a special focus on the KS method. The subsystem DFT scheme, which is an efficient alternative to the KS method, is introduced in Chapter 3.

Chapter 4 shows that the effective-embedding potentials produced by the currently available kinetic-energy density functionals for use in the frozen-density embedding (FDE) method, a subsystem formulation of DFT, appear to be too repulsive to properly describe the coordination bonding found in the noble gas atom bound to the gold fluoride molecule. The use of increasing orders of the reduced

density gradient s in the expression of generalised gradient approximation (GGA) kinetic-energy functionals does not necessarily yield improved functionals as can be seen from the bad performance of the PBE4 functional. This could, however, be an effect of the chosen parametrization that was developed for a quite different purpose and bonding situation. The performance of a given GGA functional varies significantly depending on the location on the bond axis. For example, the density obtained with the PBE3 functional follows the KS trend closely near the Ar and F atoms whereas in the bonding region and near the Au atom it fails.

The general conclusion is that for systems like NgAuF (Ng=Ar, Kr, Xe), none of the enhancement functions used in the approximate kinetic-energy functionals is yet able to describe the weak covalent bond adequately. The failure of GGA kinetic-energy functionals even for weak covalent bonds raises the question as to whether the "conjointness" approach as applied in previous work is suitable to derive functionals that can describe the stronger interactions which are of interest in many applications of subsystem methods.[134] It could be interesting to approximate the kinetic-energy component of the embedding potential directly in a non-decomposable fashion[SMB4],[84] also in regions in which the frozen and active densities show significant overlap. An alternative is to introduce approximations that not only locally depend on the electron densities of the subsystems, but that also depend on the KS orbitals of the subsystems. This would make the method similar to a pseudopotential approach in which the orbital information is also used to model the effect of the frozen density.

Chapter 5 includes using the potential reconstruction scheme presented in Section 5.1, and builds up on the work of Fux *et al.*[86] by investigating the validity of the freeze-and-thaw scheme in conjunction with the potential reconstruction for a given electron density. Two model systems with covalent bonding: LiH and Li_2, partitioned into Li^+ and H^-; Li^+ and Li^- subsystems; respectively were used. The advantage of such a partitioning is that it reduces the negativity of $\rho^{(1)}(\boldsymbol{r})$ and, hence, is a great attempt to fulfil the first condition described in Subsection 5.2.1. Furthermore, it provides two strongly different starting points for these simple model systems for which one can analyse the effect of freeze-and-thaw cycles.

The potential reconstruction calculations were started from the the superposition of the densities obtained by KS calculations on the isolated fragments (s.o.f.) as well as from FDE(s)(6) densities obtained with TF9W. To assess the quality of the reconstructed electron densities, the dipole moment difference for the total system $\Delta\mu$ (Eq. (4.15)), and the integrated difference density, Δ_{abs} (Eq. (5.5)) were used.

For both LiH and Li_2, potential reconstruction scheme started with the s.o.f.

densities improves the quality of the electron density upon TF9W as well as s.o.f.. This improvement is consistently present in successive freeze-and-thaw cycles. However, from cycle 8 to cycle 10, the difference between successive cycles gets smaller (1.0E-4 and 4.0E-04 for LiH and Li_2; respectively). This trend is also present in the negative densities.

On the other hand, it is shown that for both LiH and Li_2, potential reconstruction scheme starting from TF9W densities results in a higher quality of the electron density, with respect to that of starting from s.o.f., already without freeze-and-thaw cycles. In the case of the LiH molecule, it is noteworthy that these TF9W starting densities have a larger Δ_{abs} than that of s.of. but a lower negative density for both subsystems. The improvement obtained, therefore, shows the importance of negative density in the starting point.

For LiH, the freeze-and-thaw cycles result in an improvement until the fifth cycle, then the quality of reconstructed densities continuously decreases. For Li_2, already after the first cycle, the improvement stops. Remarkably, the electron density of the Li^- fragment obtained from this first cycle, is positive everywhere in space. Given that the electron density is v_s-representable, the next freeze-and-thaw cycle, i.e. the second cycle, should in principle result in the exact density. However, the second cycle does not improve upon the result of the first one but instead worsens it. The reason for this phenomena certainly deserves attention and is subject to future research. Identifying the causes for this problem can help improve the current reconstruction scheme.

Furthermore, for the 2-electron H^- fragment, given that the system is v_s-representable, the exact potential was calculated, which was compared with that obtained from the reconstruction scheme. It is shown that the reconstructed potentials capture the qualitative trend, however, fail to obtain the correct magnitude.

Chapter 6 is devoted to propose density functional theory including dispersion correction as approximated by the frozen-density embedding method (FDE-D) as an efficient method to calculate interaction energies between biomolecular fragments. It is shown that for selected test cases using the dispersion correction of Grimme[167] in conjunction with the BLYP XC functional and the TW02 kinetic-energy functional indicate that sufficient accuracy can be reached at a significantly lower cost compared to supermolecular DFT calculations.

In order to analyse the bonding between the fragments in chemically meaningful terms, a bond energy analysis is proposed, inspired by the KS bond formation analysis[28] but allowing for nonorthogonality of orbitals that belong to different subsystems. One can thereby write the bond energy in terms of promotion en-

ergies of individual subsystems and an explicit interaction energy between these (Eq. 6.13). This analysis provides a new view on the effect of Pauli repulsion and has a unique feature that it is done in terms of interacting fragment densities that sum to the full molecular density.

For all the molecules tested in this work, it is shown that the electrostatic attraction becomes more attractive in the course of the FT process which confirms the intuitive picture of FT providing mutual polarization.

Chapter 7 presents an explanation for the anomalous fluorescence behavior of isoflavones. Isoflavones are known to show fluorescence with intensities that depend strongly on the solvent properties and exhibit Stokes' shifts as large as 1.4 eV. While some of this behaviour can be explained by (excited state) deprotonation, this mechanism does not apply for all isoflavones.

Second-order approximate coupled cluster (CC2) calculations reproduce well the experimentally observed absorption maxima. The lowest excited singlet states are characterized by a planar arrangement of the chromone and phenyl rings with an elongated carbonyl bond. For these states charge density difference plots show major changes in this carbonyl region, with the C4-C3 of the chromone ring acquiring double bond character. The resulting conjugation of the two aromatic rings inducing a rotation towards coplanarity could explain the red-shifted emission. Since the negative charge acquired by the carbonyl oxygen in the excited state can be stabilized by an H-bonding solvent, but not by an aprotic solvent such as acetonitrile, this may also provide an explanation for the difference in fluorescence observed in the two solvents. The calculated Stokes' shift depends on the computational method employed in the structure optimization: the CC2 optimization yielded a larger carbonyl stretch and larger Stokes' shift than found with the time-dependent density functional (PBE0) structures that we had to employ in the large microsolvation models.

Chapter 9

Samenvatting

Subsysteem dichtheidsfunctionaaltheorie (DFT) is een efficiënt alternatief voor de conventionele Kohn-Sham (KS) methode omdat het een nauwkeurige beschrijving geeft tussen het subsysteem van belang en de omgeving met een kortere rekentijd. Deze interactie wordt gerepresenteerd door een effectieve potentiaal, die bestaat uit het electrostatische potentiaal van de omgeving, een exchange-correlatie component en een kinetische-energie component. De formulering van subsyteem DFT is, in tegenstelling tot veel andere embedding schema's in theoretische chemie, in principe exact. Dankzij deze genoemde voordelen is subsysteem DFT een schema met goede potentie om de grote systemen, waar het evenwicht tussen efficiëntie een nauwkeurigheid van belang zijn, succesvol te bestuderen.

Dit proefschrift evalueert verder de huidige gebreken, en de redenen daarvoor, van subsysteem DFT door een uitgebreide analyse van het effectieve embedding potentiaal. De aandacht ligt bij het gedrag van het effectieve embedding potentiaal in wiens aanwezigheid de electronendichtheid wordt bepaald van het systeem van belang. Het evalueert verder de toepassing van subsysteem DFT.

Hoofdstuk 2 geeft een review van DFT met een focus op de KS methode. Hoofdstuk 3 introduceert het sub-systeem DFT schema, welke een efficiënt alternatief voor de KS methode is.

Hoofdstuk 4 laat zien dat de effectieve embedding potentiaal, geproduceerd door op dit moment beschikbare kinetische-energie dichtheidsfunctionalen voor het gebruik in de bevroren dichtheid embedding (FDE) methode, te repulsief lijkt te zijn om de coördinatiebinding tussen het edelgas atoom en goud-fluoride molecuul correct te beschrijven. Het gebruik van toenemende ordes van de gere-

duceerd dichtheidsgradient s in de expressie van GGA kinetische-energie functionalen leidt niet per se tot verbeterde functionalen zoals uit de slechte prestatie van PBE4 blijkt. Maar dat kan ook een gevolg zijn van de gekozen parametrizatie zijn die ontwikkeld was voor een ander doel en bindingssituatie. De prestatie van een bepaald generalised gradient approximation (GGA) functionaal verandert significant afhankelijk van de locatie van de bindingsas. Bijvoorbeeld, de dichtheid verkregen door het gebruik van de PBE3 functionaal volgt de KS trend dichtbij het gebied van de Ar en F atomen. Aan de andere kant, zowel in de bindingsregio als dichtbij de Au atoom mislukt het.

De algemene conclusie is dat voor systemen zoals NgAuF (Ng=Ar, Kr, Xe), geen van de enhancement functies gebruikt in de benaderde kinetische-energie functionalen in staat is om de zwakke covalente binding adequaat te beschrijven. Dit falen van GGA kinetische-energie functionalen, zelfs voor zwakke covalente bindingen, roept de vraag op of de "conjointness" aanpak geschikt is om de functionalen te af te leiden die de sterkere interacties kunnen beschrijven, die interessant voor verschillende toepassingen van sub-systeem methodes zijn.[134] Het zou interessant zijn om de kinetische-component van het embedding potentiaal direct te benaderen via een ondeelbare manier[SMB4] ,[84] ook in de regio's waar de bevroren en actieve dichtheden een significante overlap hebben. Een alternatief is om benaderingen te introduceren die zowel lokaal afhankelijk van de electronendichtheden van de sub-systemen als van de KS orbitalen van de sub-systemen zijn. Dit zou de methode vergelijkbaar maken met een pseudo-potentiaal aanpak, waarin de orbitaalinformatie ook gebruikt wordt om het effect van de bevroren dichtheid te modelleren.

Hoofdstuk 5 omvat het gebruik van de potentiaal reconstructie schema, die gepresenteerd is in Sectie 5.1, en bouwt verder op het werk van Fux *et al.*[86] door de validiteit van het freeze-and-thaw schema te onderzoeken in combinatie met de potentiaal reconstructie. Twee modelsystemen met covalente binding waren gebruikt: LiH en Li_2, opgedeeld in respectievelijk Li^+ en H^-; Li^+ en Li^- subsystemen. Het voordeel van een dergelijke opdeling is dat het de negatieve waarde van $\rho^1_{\mathbf{r}}$ reduceert. Daarom is het een goede poging om aan de eerste conditie, beschreven in Subsectie 5.2.1, te voldoen. Bovendien voorziet het twee fors andere beginpunten voor deze simpele modelsystemen voor welke het effect van freeze-and-thaw cycles geanalyseerd kan worden.

De potentiaal herconstructie berekeningen waren begonnen van de superpositie van de dichtheden verkregen uit KS berekeningen aan de gesoleerde fragmenten (s.o.f.) en van FDE(s)(6) dichtheiden verkregen uit TF9W. Om een beoordeling van gereconstrueerde dichtheden te maken, het verschil van dipoolmo-

ment van het totale systeem $\Delta\mu$ (Eq. (4.15)), en de geïntegreerde verschillen-dichtheid, Δ_{abs} (Eq. (5.5)) waren gebruikt.

Voor zowel LiH en Li_2 was de potentiaal herconstructie schema begonnen met de s.o.f. dichtheden die de kwaliteit van de electronendichtheid ten opzichte van TF9W en s.o.f. verbetert. Deze verbetering is consistent aanwezig in de opeenvolgende freeze-and-thaw cycles. Maar vanaf cycle 8 tot en met cycle 10 wordt het verschil tussen opeenvolgende cycles kleiner (1.0E-4 and 4.0E-04 for LiH and Li_2; respectievelijk). Deze trend is ook te zien in de negatieve dichtheiden.

Aan de andere kant wordt er laten zien dat voor zowel LiH en Li_2, als het potentiaal reconstructie schema van de TF9W dichtheiden begint, het al zonder freeze-and-thaw cycles tot een hogere kwaliteit van de electronendichtheid leidt ten opzichte van waar het vanaf s.o.f. begint. Voor Li_2 stopt de verbetering al nadat de eerste cycle afgerond is. Opvallend is dat de electronendichtheid van het Li^- fragment gekregen van de eerste cycle overal in de ruimte positief is. Mits de electronendichtheid v_s gerepresenteerd kan worden, moet de volgende freeze-and-thaw cycle, i.e. de tweede cycle, leiden tot de exacte dichtheid. Maar de tweede cycle verslechtert het resultaat van de eerste cycle in plaats van het te verbeteren. De oorzaak voor dit fenomeen verdient meer aandacht en is het onderwerp voor toekomstig onderzoek. De identificatie van de oorzaken voor dit probleem kan helpen met de verbetering van het huidige potentiaal reconstructie schema.

Verder was voor het 2-electron H^- fragment, mits het systeem v_s gerepresenteerd kan worden, de exacte potentiaal berekend, die vergeleken werd met degene die geproduceerd was door het potentiaal herconstructie schema. Dit laat zien dat het de gereconstrueerde potentialen gelukt is de kwalitatieve trend te beschrijven maar dat het verkrijgen van de juiste grootte mislukt is.

Hoofdstuk 6 stelt DFT inclusief de dispersie correctie, benaderd door de methode van FDE, voor als een efficiënte methode om de interactie energieën tussen biomoleculaire fragmenten te berekenen. Voor gekozen gevallen, het gebruik van de dispersie correctie van Grimme,[167] in samenwerking met de BLYP XC functionaal en de TW02 kinetische-energie functionaal, geeft aan dat voldoende nauwkeurigheid bereikt kan worden op een significant meer efficiënte manier ten opzichte van supermolecular DFT berekeningen.

Om de binding tussen de fragmenten op een chemisch nuttige manier te analyseren, een bindingsenergie is voorgesteld, welke geïnspireerd bij de KS bindingsenergie decompositie-analyse[28] is, maar het laat nonorthogonaliteit van orbitalen van verschillende subsystemen toe. Men kan dus de bindingsenergie herschrijven gerelateerd aan de promotie energieën van individuele subsystemen en

aan een expliciet interactie energie tussen deze (Eq. 6.13). Deze analyse geeft een nieuwe blik naar het effect van Pauli repulsie en heeft een uniek kenmerk dat het gedaan wordt met dichtheiden van interagerende fragmenten die tot de volledige moleculaire dichtheid optellen.

Voor alle geteste moleculen in deze studie wordt de elektrostatische attractie meer attractief gedurende de freeze-and-thaw (FT) cycli. Dat bevestigt het intuïtieve beeld van FT cycli dat ze een wederzijdse polarisatie geven.

Hoofdstuk 7 presenteert een verklaring voor het abnormale fluorescentie gedrag van isoflavonen. Isoflavonen zijn bekend door het vertonen van fluorescentie met intensiteiten die sterk afhankelijk zijn van de eigenschappen van het oplosmiddel zijn en ze tonen Stoke's shifts zo groot als 1.4 eV. Voor een deel kan dit gedrag verklaard worden door geëexciteerde toestand deprotonatie, maar dit mechanisme geldt niet voor alle isoflavonen.

Tweede-orde benaderde coupled cluster (CC2) berekeningen reproduceren de experimentele absorptiespectra goed. De laagste geëxciteerde singlet toestanden zijn gekarakteriseerd door een vlakke ordening van de chromone en phenyl ringen met een verlengde carbonyl binding. Voor deze toestanden laten de dichtheidsverschil grafieken significante veranderingen zien in deze carbonyl regio en dat de C4-C3 van de chromone ring een dubbele binding karakter krijgt. De resulterende conjugatie van de twee aromatische ringen die een rotatie richting vlakheid induceert kan de rood-gestifte emissie uitleggen. Omdat de negatieve lading gekregen door het carbonyl zuurstof in geëxciteerde toestand gestabiliseerd kan worden door een waterstofverbonden oplosmiddel, maar niet door een carbonyl zuurstof, kan dit geobserveerde verschil in fluorescentie in twee oplosmiddelen uitleggen. De berekende Stokes's shifts zijn afhankelijk van de rekenmethode die in de struktuuroptimalisatie toegepast is: CC2 optimalisatie geeft een grotere uitrekking van carbonyl en een grotere Stokes' shift in vergelijking met degene gevonden in structuren verkregen door de tijdsafhankelijke dichtheidsfunctionaal (PBE0) die in de microsolvatie modellen toegepast was.

Chapter 10

Dankwoord

Aan iedereen die dank verdient, dank.

List of Publications

[SMB1]S. M. Beyhan, A. W. Götz, , F. Ariese, L. Visscher, and C. Gooijer, J. Phys. Chem. **115**, 1493 (2010).

[SMB2]A. W. Gotz, S. M. Beyhan, and L. Visscher, J. Chem. Theory Comput. **5**, 3161 (2009).

[SMB3]S. M. Beyhan, A. W. Götz, and L. Visscher, Submitted to J. Chem. Phys. **xxx**, xxx (2012).

[SMB4]C. R. Jacob, S. M. Beyhan, and L. Visscher, J. Chem. Phys. **126**, 234116 (2007).

[SMB5]S. M. Beyhan, A. W. Götz, C. R. Jacob, and L. Visscher, J. Chem. Phys. **132**, 044114 (2010).

[SMB6]C. R. Jacob, S. M. Beyhan, R. E. Bulo, A. S. P. Gomes, A. W. Götz, K. Kiewisch, J. Sikkema, and L. Visscher, Phys. Chem. Chem. Phys. **32**, 2328 (2011).

Bibliography

[1] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic Structure Theory*, John Wiley & Sons, Chichester, 2000.

[2] W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, Wiley-WCH, Weinheim, 2nd edition, 2001.

[3] F. Maseras and K. Morokuma, J. Comput. Chem **16**, 1170 (1995).

[4] K. Kitaura, Chem. Phys. Lett. **313**, 701 (1999).

[5] D. W. Zhang and J. Z. H. Zhang, J. Chem. Phys. **119**, 3599 (2003).

[6] R. P. A. Bettens and A. M. Lee, J. Phys. Chem. A **110**, 8777 (2006).

[7] N. Vaidehi, T. A. Wesolowski, and A. Warshel, J. Chem. Phys. **97**, 4264 (1992).

[8] A. Warshel and M. Levitt, J. Mol. Biol. **103**, 227 (1976).

[9] U. Ryde, Curr. Opin. Chem. Biol. **7**, 136 (2003).

[10] M. Peräkylä, *Quantum Medicinal Chemistry*, pages 157–176, Wiley-WCH, Weinheim, 2003.

[11] F. Perruccio, L. Ridder, and A. J. Mulholland, *Quantum Medicinal Chemistry*, pages 177–198, Wiley-WCH, Weinheim, 2003.

[12] J. L. Rivail, *Computational Medicinal Chemistry for Drug Discovery*, pages 119–131, Dekker, New York, 2004.

[13] R. A. Friesner, Drug Discovery Today Technol. **1**, 253 (2004).

[14] A. J. Mulholland, Drug. Discov. Today **10**, 1393 (2005).

[15] R. A. Friesner and V. Guallar, Annu. Rev. Phys. Chem. **56**, 389 (2005).

[16] H. M. Senn and W. Thiel, Curr. Opin. Chem. Biol. **11**, 182 (2007).

[17] H. Lin and D. G. Truhlar, Theoret. Chem. Acc. **117**, 185 (2007).

[18] H. Hu and W. Yang, Annu. Rev. Phys. Chem **59**, 573 (2008).

[19] H. M. Senn and W. Thiel, Angew. Chem. Int. Edit. **48**, 1198 (2009).

[20] K. E. Ranaghan and A. J. Mulholland, Int. Rev. Phys. Chem. **29**, 65 (2010).

[21] T. A. Wesolowski and A. Warshel, J. Phys. Chem. **97**, 8050 (1993).

[22] S. Humbel, S. Sieber, and K. Morokuma, J. Chem. Phys. **105**, 1959 (1996).

[23] M. Svensson, S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber, and K. Morokuma, J. Phys. Chem. **100**, 19357 (1996).

[24] S. Borini, D. Maynau, and S. Evangelisti, J. Comput. Chem **26**, 1042 (2005).

[25] T. M. Henderson, J. Chem. Phys. **125**, 014105 (2006).

[26] P. Cortona, Phys. Rev. B **44**, 8454 (1991).

[27] C. R. Jacob, *Frozen Density Embedding*, PhD thesis, Vrije Universiteit, Amsterdam, 2007.

[28] F. M. Bickelhaupt and E. J. Baerends, Rev. Comput. Chem. **15**, 1 (2000).

[29] I. N. Levine, *Quantum Mechanics*, Prentice Hall, New Jersey, 5th edition, 1999.

[30] W. Kutzelnigg, *Einführung in die theoretische Chemie, Band 1 und Band 2*, Verlag Chemie, Weinheim, 1978.

[31] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry*, Dover Publications, Mineola, N.Y., 1996.

[32] G. D. Purvis and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).

[33] U. Schollwock, Rev. Mod. Phys. **77**, 259 (2005).

[34] P. Hohenberg and W. Kohn, Phys. Rev. B **136**, B864 (1964).

[35] W. Kohn and L. J. Sham, Phys. Rev. **140**, 1133 (1965).

[36] R. G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules*, Oxford University Press, Oxford, 1989.

[37] M. Levy, Phys. Rev. A **26**, 1200 (1982).

[38] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[39] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).

[40] J. P. Perdew, Unified theory of exchange and correlation beyond the local density approximation, in *Electronic Structure of Solids*, edited by P. Ziesche and H. Eschrig, pages 11–20, Akademie Verlag, Berlin, 1991.

[41] J. P. Perdew, Phys. Rev. B **33**, 8822 (1986).

[42] C. T. Lee, W. T. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).

[43] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, Phys. Rev. Lett. **91**, 146401 (2003).

[44] P. R. T. Schipper, O. V. Gritsenko, S. J. A. van Gisbergen, and E. J. Baerends, J. Chem. Phys. **112**, 1344 (2000).

[45] M. Grüning, *Density functional theory with improved gradient and orbital dependent functionals*, PhD thesis, Vrije Universiteit, Amsterdam, 2003.

[46] M. Gruning, O. V. Gritsenko, and E. J. Baerends, J. Chem. Phys. **118**, 7183 (2003).

[47] A. Gorling, J. Chem. Phys. **123**, 062203 (2005).

[48] J. Černy and P. Hobza, Phys. Chem. Chem. Phys. **7**, 1624 (2005).

[49] M. J. Allen and D. J. Tozer, J. Chem. Phys. **117**, 11113 (2002).

[50] P. Hobza, J. Šponer, and T. Reschel, J. Comput. Chem **16**, 1315 (1995).

[51] S. Kristiyan and P. Pulay, Chem. Phys. Lett. **229**, 175 (1994).

[52] S. Grimme, J. Antony, T. Schwabe, and C. Mueck-Lichtenfeld, Org. Biomol. Chem. **5**, 741 (2007).

[53] J. Gräfenstein and D. Cremer, J. Chem. Phys. **130** (2009).

[54] E. R. Johnson, I. D. Mackie, and G. A. DiLabio, J. .Phys. Org. Chem. **22**, 1127 (2009).

[55] T. Sato and H. Nakai, J. Chem. Phys. **131** (2009).

[56] Y. A. Wang and E. A. Carter, *Theoretical Methods in Condensed Phase Chemistry*, chapter 5, pages 117–184, Kluwer, Dordrecht, 2000.

[57] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, Phys. Rev. Lett. **49**, 1691 (1982).

[58] O. V. Gritsenko and L. Visscher, Phys. Rev. A **82**, 032519 (2010).

[59] T. A. Wesolowski and A. Warshel, J. Phys. Chem. **97**, 8050 (1993).

[60] T. A. Wesolowski, One-electron equations for embedded electron density: challenge for theory and practical payoffs in multilevel modelling of complex polyatomic systems, in *Computational Chemistry: Reviews of Current Trends*, edited by J. Leszczynski, volume 10, World Scientific, Singapore, 2006.

[61] T. Wesolowski and A. Warshel, J. Phys. Chem. **98**, 5183 (1994).

[62] J. Neugebauer, M. J. Louwerse, E. J. Baerends, and T. A. Wesolowski, J. Chem. Phys. **122**, 094115 (2005).

[63] J. Neugebauer, C. R. Jacob, T. A. Wesolowski, and E. J. Baerends, J. Phys. Chem. A **109**, 7805 (2005).

[64] J. Neugebauer, M. J. Louwerse, P. Belanzoni, T. A. Wesolowski, and E. J. Baerends, J. Chem. Phys. **123**, 114101 (2005).

[65] C. R. Jacob, J. Neugebauer, L. Jensen, and L. Visscher, Phys. Chem. Chem. Phys. **8**, 2349 (2006).

[66] A. S. P. Gomes, C. R. Jacob, and L. Visscher, Phys. Chem. Chem. Phys. **10**, 5353 (2008).

[67] P. Cortona, Phys. Rev. B **44**, 8454 (1991).

[68] T. A. Wesolowski and J. Weber, Chem. Phys. Lett. **248**, 71 (1996).

[69] O. V. Gritsenko, *Recent Advances in Orbital-Free Density Functional Theory*, page 383, World Scientific, Singapore, 2010.

[70] M. Iannuzzi, B. Kirchner, and J. Hutter, Chem. Phys. Lett. **421**, 16 (2006).

[71] T. A. Wesolowski, H. Chermette, and J. Weber, J. Chem. Phys. **105**, 9182 (1996).

[72] T. A. Wesolowski, J. Chem. Phys. **106**, 8516 (1997).

[73] T. A. Wesolowski and J. Weber, Int. J. Quantum Chem. **61**, 303 (1997).

[74] G. Senatore and K. R. Subbaswamy, Phys. Rev. B **34**, 5754 (1986).

[75] R. Kevorkyants, M. Dulak, and T. A. Wesolowski, J. Chem. Phys. **124**, 024104 (2006).

[76] M. Dulak and T. A. Wesolowski, J. Mol. Model **13**, 631 (2007).

[77] M. Dulak, J. W. Kaminski, and T. A. Wesolowski, J. Chem. Theory Comput. **3**, 735 (2007).

[78] T. A. Wesolowski, Y. Ellinger, and J. Weber, J. Chem. Phys. **108**, 6078 (1998).

[79] T. A. Wesolowski, P.-Y. Morgantini, and J. Weber, J. Chem. Phys. **116**, 6411 (2002).

[80] T. A. Wesolowski and F. Tran, J. Chem. Phys. **118**, 2072 (2003).

[81] C. R. Jacob, T. A. Wesolowski, and L. Visscher, J. Chem. Phys. **123**, 174104 (2005).

[82] K. Kiewisch, G. Eickerling, M. Reiher, and J. Neugebauer, J. Chem. Phys. **128**, 044114 (2008).

[83] S. Fux, K. Kiewisch, C. R. Jacob, J. Neugebauer, and M. Reiher, Chem. Phys. Lett. **461**, 353 (2008).

[84] J. M. Garcia-Lastra, J. W. Kaminski, and T. A. Wesolowski, J. Chem. Phys. **129**, 074107 (2008).

[85] H. Lee, C. Lee, and R. G. Parr, Phys. Rev. A **44**, 768 (1991).

[86] S. Fux, C. R. Jacob, J. Neugebauer, L. Visscher, and M. Reiher, J. Chem. Phys. **132**, 164101 (2010).

[87] C. R. Jacob and L. Visscher, J. Chem. Phys. **128**, 155102 (2008).

[88] T. A. Wesolowski, Chem. Phys. Lett. **311**, 87 (1999).

[89] R. E. Bulo, C. R. Jacob, and L. Visscher, J. Phys. Chem. A **112**, 2640 (2008).

[90] M. E. Casida and T. A. Wesolowski, Int. J. Quantum Chem. **96**, 577 (2004).

[91] T. A. Wesolowski, J. Am. Chem. Soc. **126**, 11444 (2004), PMID: 15366883.

[92] J. Neugebauer and E. J. Baerends, J. Phys. Chem. A **110**, 8786 (2006), PMID: 16836441.

[93] J. Neugebauer, J. Chem. Phys. **126**, 134116 (2007).

[94] N. Govind, Y. A. Wang, A. J. R. da Silva, and E. A. Carter, Chem. Phys. Lett. **295**, 129 (1998).

[95] N. G. N, Y. A. Wang, and E. A. Carter, J. Chem. Phys. **110**, 7677 (1999).

[96] T. Kluner, N. Govind, Y. A. Wang, and E. A. Carter, Phys. Rev. Lett. **86**, 5954 (2001).

[97] T. Kluner, N. Govind, Y. A. Wang, and E. A. Carter, J. Chem. Phys. **116**, 42 (2002).

[98] P. Huang and E. A. Carter, J. Chem. Phys. **125**, 084102 (2006).

[99] M. Zbiri, M. Atanasov, C. Daul, J. M. Garcia-Lastra, and T. A. Wesolowski, Chem. Phys. Lett. **397**, 441 (2004).

[100] J. Neugebauer, J. Phys. Chem. B **112**, 2207 (2008).

[101] M. Dulak and T. A. Wesolowski, Int. J. Quantum Chem. **101**, 543 (2005).

[102] T. A. Wesolowski, Y. Ellinger, and J. Weber, J. Chem. Phys. **108**, 6078 (1998).

[103] T. A. Wesolowski, J. Chem. Phys. **106**, 8516 (1997).

[104] F. Tran, J. Weber, and T. A. Wesolowski, Helv. Chim. Acta **84**, 1489 (2001).

[105] V. V. Karasiev, S. B. Trickey, and F. E. Harris, J. Comp.-Aid. Mat. Des. **14**, 111 (2006).

[106] S. A. Cooke and M. C. L. Gerry, J. Am. Chem. Soc. **126**, 17000 (2004).

[107] L. Belpassi, I. Infante, F. Tarantelli, and L. Visscher, J. Am. Chem. Soc. **130**, 1048 (2008).

[108] L. H. Thomas, Proc. Camb. Phys. **23**, 542 (1927).

[109] E. Fermi, Atti. Accad. Nazl. Lincei **6**, 602 (1927).

[110] E. H. Lieb, Rev. Mod. Phys. **53**, 603 (1981).

[111] E. Teller, Rev. Mod. Phys. **34**, 627 (1962).

[112] N. L. Balàzs, Phys. Rev. **156**, 42 (1967).

[113] E. H. Lieb and B. Simon, Phys. Rev. Lett. **31**, 681 (1973).

[114] E. H. Lieb and B. Simon, Adv. Math **23**, 22 (1977).

[115] C. F. von Weizsäcker, Z. f. Phys. **96**, 431 (1935).

[116] A. Lembarki and H. Chermette, Phys. Rev. A **50**, 5328 (1994).

[117] D. J. Lacks and R. G. Gordon, J. Chem. Phys. **100**, 4446 (1994).

[118] P. Fuentealba and O. Reyes, Chem. Phys. Lett. **232**, 31 (1995).

[119] P. Fuentealba, J. Mol. Struct. **390**, 1 (1997).

[120] F. Tran and T. A. Wesolowski, Int. J. Quantum Chem. **89**, 441 (2002).

[121] J. P. Perdew, Phys. Lett. A **165**, 79 (1992).

[122] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[123] C. Adamo and V. Barone, J. Chem. Phys. **116**, 5933 (2002).

[124] C. R. Jacob, J. Neugebauer, and L. Visscher, J. Comput. Chem **29**, 1011 (2008).

[125] G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders, and T. Ziegler, J. Comput. Chem **22**, 931 (2001).

[126] ADF2008.01, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, http://www.scm.com, accessed on May 2008.

[127] E. van Lenthe, E. J. Baerends, and J. G. Snijders, J. Chem. Phys. **99**, 4597 (1993).

[128] E. van Lenthe, E. J. Baerends, and J. G. Snijders, J. Chem. Phys. **101**, 9783 (1994).

[129] G. van Rossum and J. de Boer, *Interactively Testing Remote Servers Using the Python Programming Language*, volume 4, CWI Quarterly, Amsterdam, 1991.

[130] C. J. Evans, D. S. Rubinoff, and M. C. L. Gerry, Phys. Chem. Chem. Phys. **2**, 3943 (2000).

[131] J. M. Thomas, N. R. Walker, S. A. Cooke, and M. C. L. Gerry, J. Am. Chem. Soc. **126**, 1235 (2004).

[132] B. P. Tripathi, R. K. Laloraya, and S. L. Srivasta, Phys. Rev. A **6**, 850 (1972).

[133] Y. A. Bernard, M. Dulak, J. W. Kaminski, and T. A. Wesolowski, J. Phys. A-Math. Theor. **41**, 055302 (2008).

[134] C. R. Jacob, *Recent Advances in Orbital-Free Density Functional Theory*, page in press, World Scientific, Singapore, 2010.

[135] O. V. Gritsenko, P. R. T. Schipper, and E. J. Baerends, Chem. Phys. Lett. **302**, 199 (1999).

[136] O. V. Gritsenko, P. R. T. Schipper, and E. J. Baerends, Int. J. Quantum Chem. **76**, 407 (2000).

[137] T. W. Keal and D. J. Tozer, J. Chem. Phys. **119**, 3015 (2003).

[138] T. W. Keal and D. J. Tozer, J. Chem. Phys. **121**, 5654 (2004).

[139] A. Savin and T. A. Wesolowski, *Advances in the Theory of Atomic and Molecular Systems*, pages 311–326, Springer, Dordrecht, 2009.

[140] Q. Wu and W. Yang, J. Chem. Phys. **118**, 2498 (2003).

[141] F. A. Bulat, T. Heaton-Burgess, A. J. Cohen, and W. Yang, J. Chem. Phys. **127**, 174101 (2007).

[142] S. Hirata, S. Ivanov, I. Grabowski, R. J. Bartlett, K. Burke, and J. D. Talman, J. Chem. Phys. **115**, 1635 (2001).

[143] V. N. Staroverov, G. E. Scuseria, and E. R. Davidson, J. Chem. Phys. **124**, 141103 (2006).

[144] T. Heaton-Burgess, F. A. Bulat, and W. Yang, Phys. Rev. Lett. **98**, 256401 (2007).

[145] A. Hesselmann, A. W. Gotz, F. D. Sala, and A. Gorling, J. Chem. Phys. **127**, 054102 (2007).

[146] C. Kollmar and M. Filatov, J. Chem. Phys. **127**, 114104 (2007).

[147] Wolfram research inc, mathematica version 6.0.3, 2005.

[148] P. Hobza, R. Zahradnik, and K. Muller-Dethlefs, Collect. Czech. Chem. Commun. **71**, 443 (2006).

[149] B. Brutschy and P. Hobza, Chem. Rev. **100**, 3861 (2000).

[150] S. Tsuzuki and T. Uchimaru, Curr. Org. Chem. **10**, 745 (2006).

[151] S. Tsuzuki, Struct. Bonding , 149 (2005).

[152] M. O. Sinnokrot, E. F. Valeev, and C. D. Sherrill, J. Am. Chem. Soc. **124**, 10887 (2002).

[153] M. O. Sinnokrot and C. D. Sherrill, The Journal of Physical Chemistry A **108**, 10200 (2004).

[154] P. Hobza and J. Šponer, J. Am. Chem. Soc. **124**, 11802 (2002).

[155] M. L. Leininger, I. M. B. Nielsen, M. E. Colvin, and C. L. Janssen, J. Phys. Chem. A **106**, 3850 (2002).

[156] R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, Oxford, 1989.

[157] Y. Andersson, D. C. Langreth, and B. I. Lundqvist, Phys. Rev. Lett. **76**, 102 (1996).

[158] D. C. Langreth, M. Dion, H. Rydberg, E. Schroder, and B. I. L. P Hyldgaard, Int. J. Quantum Chem. **101**, 599 (2005).

[159] C. D. Sherrill, Acc. Chem. Res. (2012).

[160] G. Jansen and A. Heßselmann, J. Phys. Chem. A **105**, 11156 (2001).

[161] M. Elstner, P. Hobza, T. Frauenheim, S. Suhai, and E. Kaxiras, J. Chem. Phys. **114**, 5149 (2001).

[162] S. Grimme, J. Comput. Chem **25**, 1463 (2004).

[163] P. Jurečka, J. Cerny, P. Hobza, and D. R. Salahub, J. Comput. Chem **28**, 555 (2007).

[164] T. Sato, T. Tsuneda, and K. Hirao, Molec. Phys. **103**, 1151 (2005).

[165] O. A. von Lilienfeld, I. Tavernelli, U. Rothlisberger, and D. Sebastiani, Phys. Rev. Lett. **93**, 153004 (2004).

[166] Y. Y. Sun, Y.-H. Kim, K. Lee, and S. B. Zhang, J. Chem. Phys. **129**, 154102 (2008).

[167] S. Grimme, J. Comput. Chem **27**, 1787 (2006).

[168] A. Tkatchenko and M. Scheffler, Phys. Rev. Lett. **102**, 073005 (2009).

[169] E. Bodo, F. A. Gianturco, R. Martinazzo, F. Paesani, and M. Raimondi, J. Chem. Phys. **113**, 11071 (2000).

[170] X. Wu, M. C. Vargas, S. Nayak, V. Lotrich, and G. Scoles, J. Chem. Phys. **115**, 8748 (2001).

[171] U. Zimmerli, M. Parrinello, and P. Koumoutsakos, J. Chem. Phys. **120**, 2693 (2004).

[172] G. Murdachaew, S. de Gironcoli, and G. Scoles, J. Phys. Chem. A **112**, 9993 (2008).

[173] S. N. Steinmann, G. Csonka, and C. Corminboeuf, J. Chem. Theory Comput. **5**, 2950 (2009).

[174] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. **132**, 154104 (2010).

[175] J. Antony and S. Grimme, Phys. Chem. Chem. Phys. **8**, 5287 (2006).

[176] M. Dulak and T. A. Wesolowski, J. Mol. Model. **13**, 631 (2007).

[177] H. C. Chow and S. H. Vosko, Can. J. Phys. **58**, 497 (1980).

[178] L. H. Thomas, Proc. Camb. Phil. Soc. **23**, 542 (1927).

[179] E. Fermi, Rend. Accad. Lincei **6**, 602 (1927).

[180] J. P. Perdew, J. Chevary, S. Vosko, K. A. Jackson, M. R. Pederson, D. Singh, and C. Fiolhais, Phys. Rev. B **46**, 6671 (1992).

[181] A. Lembarki and H. Chermette, Phys. Rev. A **50**, 5328 (1994).

[182] K. Kitaura and K. Morokuma, Int. J. Quantum Chem. **10**, 325 (1976).

[183] T. Ziegler and A. Rauk, Theoret. Chim. Acta **46**, 1 (1977).

[184] D. G. Fedorov and K. Kitaura, J. Chem. Phys. **28**, 222 (2007).

[185] P. Su and H. Li, J. Chem. Phys. **131**, 014102 (2009).

[186] J. Nafziger, Q. Wu, and A. Wasserman, J. Chem. Phys. **135**, 234101 (2011).

[187] D. G. Fedorov and K. Kitaura, J. Phys. Chem. A **116**, 704 (2012).

[188] T. Nagata, D. G. Fedorov, T. Sawada, and K. Kitaura, J. Phys. Chem. A **116**, 9088 (2012).

[189] P. Su, H. Liu, and W. Wu, J. Chem. Phys. **137**, 034111 (2012).

[190] F. R. Manby, M. Stella, J. D. Goodpaster, and T. F. Miller, J. Chem. Theory Comput. **8**, 2564 (2012).

[191] Y. Zhao and D. G. Truhlar, Phys. Chem. Chem. Phys. **7**, 2701 (2005).

[192] J. Vondrášek, L. Bendová, V. Klusák, and P. Hobza, J. Am. Chem. Soc. **127**, 2615 (2005).

[193] K. Berka, P. Hobza, and J. Vondrášek, Comp. Phys. Comm. **10**, 543 (2009).

[194] ADF2009.01, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, http://www.scm.com, accessed in November 2009.

[195] C. F. Guerra, J. Snijders, G. te Velde, and E. J. Baerends, Theoret. Chem. Acc. **99**, 391 (1998).

[196] G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. F. Guerra, S. J. A. van Gisbergen, J. G. Snijders, and T. Ziegler, J. Comput. Chem **22**, 931 (2001).

[197] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).

[198] C. T. Lee, W. T. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).

[199] B. G. Johnson, P. M. W. Gill, and J. A. Pople, J. Chem. Phys. **98**, 5612 (1993).

[200] T. V. Russo, R. L. Martin, and P. J. Hay, J. Chem. Phys. **101**, 7729 (1994).

[201] S. F. Boys and F. Bernardi, Molec. Phys. **19**, 553 (1970).

[202] P. Jurečka and P. Hobza, J. Am. Chem. Soc. **125**, 15608 (2003).

[203] P. Jurečka, J. Šponer, and P. Hobza, J. Phys. Chem. B **108**, 5466 (2004).

[204] J. Šponer, P. Jurečka, and P. Hobza, J. Am. Chem. Soc. **126**, 10142 (2004).

[205] B. Halliwell, J. Rafter, and A. Jenner, Am. J. Clin. Nutr **81**, 268S (2005).

[206] H. Wiseman, P. Nutr. Soc. **58**, 139 (1999).

[207] J. B. Harborne and C. A. Williams, Phytochemistry **55**, 481 (2000).

[208] K. Shimoi, S. Masuda, B. Shen, M. Furugori, and N. Kinae, Mutat. Res-Fund. Mol. M. **350**, 153 (1996).

[209] P. Knekt, R. Jarvinen, R. Seppanen, M. Heliovaara, L. Teppo, E. Pukkala, and A. Aromaa, Am. J. Epidemiol **146**, 223 (1997).

[210] A. Monasterio, M. Urdaci, I. Pinchuk, N. Lopez-Moratalla, and J. Martinez-Irujo, Nutr. Cancer **50**, 90 (2004).

[211] J. Terao, Y. Kawai, and K. Murcita, Asia Pac. J. Clin. Nutr. **17**, 291 (2008).

[212] J. Catalan, J. C. D. Valle, C. Diaz, J. Palomar, J. L. G. D. Paz, and M. Kasha, Int. J. Quantum Chem. **72**, 421 (1999).

[213] E. Falkovskaia, P. K. Sengupta, and M. Kasha, Chem. Phys. Lett. **297**, 109 (1998).

[214] O. S. Wolfbeis, M. Leiner, P. Hochmuth, and H. Geiger, Ber. Bunsen-Ges. Phys. Chem. **88**, 759 (1984).

[215] P. T. Chou, Y. C. Chen, W. S. Yu, and Y. M. Cheng, Chem. Phys. Lett. **340**, 89 (2001).

[216] S. P. De, S. Ash, H. Bar, D. K. Bhui, S. Dalai, and A. Misra, Theochem (J. Mol. Struct.) **824**, 8 (2007).

[217] D. McMorrow and M. Kasha, J Phys Chem-Us **88**, 2235 (1984).

[218] P. K. Sengupta and M. Kasha, Chem. Phys. Lett. **68**, 382 (1979).

[219] B. J. Schwartz, L. A. Peteanu, and C. B. Harris, J Phys Chem-Us **96**, 3591 (1992).

[220] J. Toth, M. Remko, and M. Nagy, Z Naturforsch C **51**, 784 (1996).

[221] S. Ameer-Beg, S. M. Ormson, R. G. Brown, P. Matousek, M. Towrie, E. T. J. Nibbering, P. Foggi, and F. V. R. Neuwahl, J. Phys. Chem. A **105**, 3709 (2001).

[222] A. N. Bader, F. Ariese, and C. Gooijer, J. Phys. Chem. A **106**, 2844 (2002).

[223] M. Sarkar, J. G. Ray, and P. K. Sengupta, J. Photochem. Photobiol. A **95**, 157 (1996).

[224] M. Sarkar and P. K. Sengupta, J. Photochem. Photobiol. A **48**, 175 (1989).

[225] R. Schipfer, O. S. Wolfbeis, and A. Knierzinger, J. Chem. Soc., Perkin Trans. 2 , 1443 (1981).

[226] M. Itoh, K. Hasegawa, and Y. Fujiwara, J. Am. Chem. Soc. **108**, 5853 (1986).

[227] M. Sarkar and P. K. Sengupta, J. Photochem. Photobiol. A **53**, 191 (1990).

[228] C. M. Marian, Spectrochim. Acta A **73**, 1 (2009).

[229] J. R. Udgata and S. N. Naik, J. Sci. Ind. Res. India **66**, 11 (2007).

[230] O. S. Wolfbeis and R. Schipper, Photochem. Photobiol. **34**, 567 (1981).

[231] O. S. Wolfbeis, E. Fürlinger, H. C. Jha, and F. Zilliken, Z. Naturforschung **39b**, 238 (1984).

[232] E. de Rijke, H. C. Joshi, H. R. Sanderse, F. Ariese, U. A. T. Brinkman, and C. Gooijer, Analyt. Chim. Acta **468**, 3 (2002).

[233] C. L. Dunford, G. J. Smith, E. E. Swinny, and K. R. Markham, Photoch. Photobio. Sci. **2**, 611 (2003).

[234] M. E. Casida, Part 1, in *Recent Advances in Density Functional Methods*, edited by D. P. Chong, page 155, World Scientific, Singapore, 1995.

[235] E. K. U. Gross, J. F. Dobson, and M. Petersilka, Topics in current chemistry, in *Density Functional Theory*, edited by R. Nalewajski, volume 181, page 81, Springer-Verlag, Heidelberg, 1996.

[236] E. Runge and E. K. U. Gross, Phys. Rev. Lett. **52**, 997 (1984).

[237] O. Christiansen, H. Koch, and P. Jorgensen, Chem. Phys. Lett. **243**, 409 (1995).

[238] R. J. Cave, K. Burke, and E. W. Castner, J. Phys. Chem. A **106**, 9294 (2002).

[239] D. Jacquemin, E. A. Perpète, G. Scalmani, M. J. Frisch, I. Ciofini, and C. Adamo, J. Chem. Phys. **125**, 164324 (2006).

[240] D. Jacquemin, E. A. Perpéte, X. Assfeld, G. Scalmani, M. J. Frisch, and C. Adamo, Chem. Phys. Lett. **438**, 208 (2007).

[241] M. Sulpizi, P. Carloni, J. Hutter, and U. Rothlisberger, Phys. Chem. Chem. Phys. **5**, 4798 (2003).

[242] J. Preat, D. Jaquemin, and E. A. Perpéte, Chem. Phys. Lett. **415**, 20 (2005).

[243] Y. L. Wang and G. S. Wu, Acta Phys-chim Sin **23**, 1831 (2007).

[244] M. Ernzerhof and G. E. Scuseria, J. Chem. Phys. **111**, 911 (1999).

[245] C. Adamo, G. E. Scuseria, and V. Barone, J. Chem. Phys. **111**, 2889 (1999).

[246] R. Ahlrichs, M. Bär, M. Häser, H. Horn, and C. Kölmel, Chem. Phys. Lett. **162**, 165 (1989).

[247] For the current version, see http://www.turbomole.com.

[248] O. Treutler and R. Ahlrichs, J. Chem. Phys. **102**, 346 (1995).

[249] A. Schäfer, C. Huber, and R. Ahlrichs, J. Chem. Phys. **100**, 5829 (1994).

[250] R. Bauernschmitt and R. Ahlrichs, Chem. Phys. Lett. **256**, 454 (1996).

[251] F. Furche and R. Ahlrichs, J. Chem. Phys. **117**, 7433 (2002).

[252] F. Furche and D. Rappoport, Computational and theoretical chemistry, in *Computational Photochemistry*, edited by M. Olivucci, volume 16, chapter III, Elsevier, Amsterdam, 2005.

[253] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).

[254] R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).

[255] F. Weigend, Phys. Chem. Chem. Phys. **4**, 4285 (2002).

[256] C. Hättig, A. Hellweg, and A. Köhn, Phys. Chem. Chem. Phys. **8**, 1159 (2006).

[257] C. Hättig, J. Chem. Phys. **118**, 7751 (2003).

[258] C. Hättig and F. Weigend, J. Chem. Phys. **113**, 5154 (2000).

[259] C. Hättig and K. Hald, Phys. Chem. Chem. Phys. **4**, 2111 (2002).

[260] C. Hättig and A. Köhn, J. Chem. Phys. **117**, 6939 (2002).

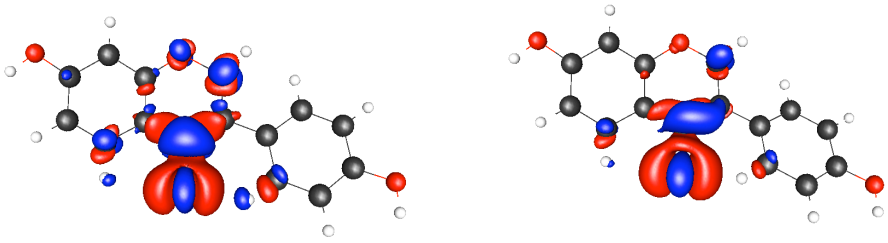
[261] C. Hättig, A. Köhn, and K. Hald, J. Chem. Phys. **116**, 5401 (2002).

[262] A. Köhn and C. Hättig, J. Chem. Phys. **119**, 5021 (2003).

[263] M. Schreiber, M. R. Silva-Junior, S. P. A. Sauer, and W. Thiel, JCP **128**, 134110 (2008).

Supporting Information

(a) PBE0/LE (S_2) (b) CC2/LE (S_1)



(c) PBE0/CT (S_1) (d) CC2/CT (S_2)

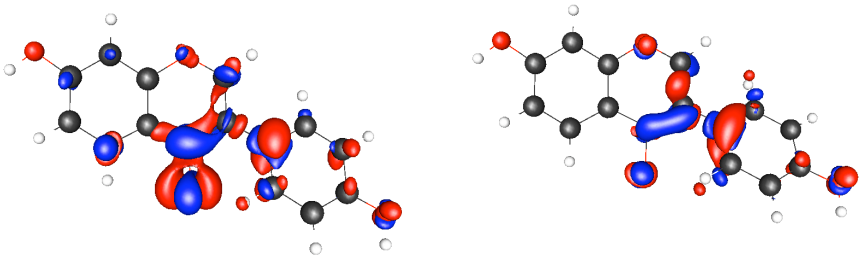


Figure 1: Gas phase charge density difference plots at the optimized geometries of the lowest two excitations (S_1 and S_2) at the PBE0/TZVP and CC2/aug-cc-pVDZ levels. CT indicates a charge transfer state whereas LE a locally excited state. Isosurface value: 0.003. Blue indicates an increase in the charge density and red a decrease.

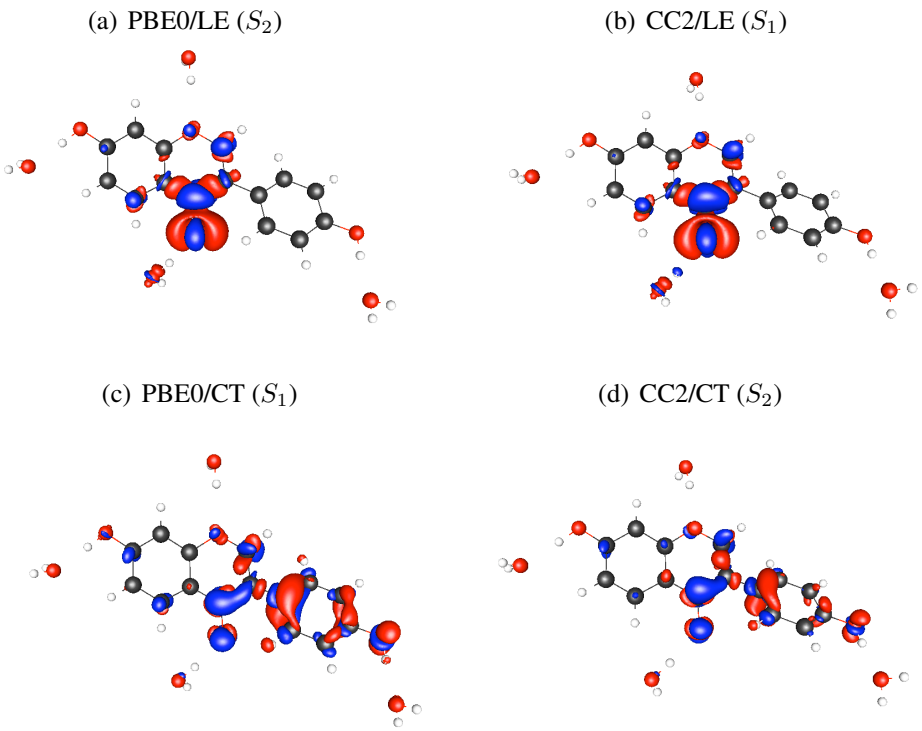
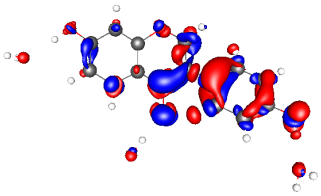
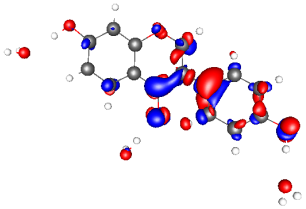


Figure 2: Water microsolvation (D+4H_2O) charge density difference plots at the ground state geometry for the lowest two excitations (S_1 and S_2) at the PBE0/TZVP and CC2/aug-cc-pVDZ levels. CT indicates a charge transfer state whereas LE a locally excited state. Isosurface value: 0.003. Blue indicates an increase in the charge density and red a decrease.

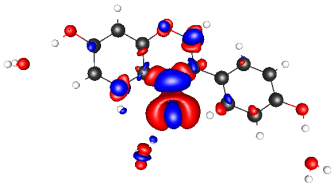
(a) PBE0/CT (S_1)



(b) CC2/CT (S_1)



(c) PBE0/LE (S_2)



(d) CC2/LE (S_2)

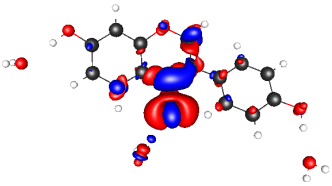


Figure 3: Water microsolvation (D+3H_2O) charge density difference plots of the lowest two excited states (S_1 and S_2) at the PBE0/TZVP optimised charge transfer (CT) (S_1) geometry at the PBE0/TZVP and CC2/aug-cc-pVDZ levels. LE indicates the locally excited state. Isosurface value: 0.003. Blue indicates an increase in the charge density and red a decrease.

Table 1: XYZ coordinates of water microsolvation (D+3H_2O) PBE0/TZVP optimized ground state, charge transfer (CT) and locally excited (LE) excited states:

```
38
Ground State
O     1.3024514   -2.3378520    0.0998070
C    -0.0208605   -2.1246334   -0.0179418
C    -0.5577677   -0.8685823   -0.1194028
C     0.2671374    0.3143501   -0.1653599
O    -0.2236150    1.4785104   -0.3192830
C     1.6878628    0.0582436   -0.0499819
C     2.1663508   -1.2601894    0.0672669
C     2.6482979    1.0756446   -0.0627781
C     4.0121564    0.7939708    0.0391420
C     4.4438427   -0.5164411    0.1492995
C     3.5030845   -1.5619066    0.1634360
H     2.3082011    2.1009038   -0.1450748
H     3.8304389   -2.5895981    0.2514167
C    -2.0122872   -0.7572640   -0.0914325
C    -2.6311025    0.3921914    0.4776618
C    -3.9889095    0.5371073    0.4784046
C    -4.8023721   -0.4741439   -0.0777900
C    -4.2099475   -1.6418257   -0.6180693
C    -2.8511569   -1.7725596   -0.6196803
H    -1.9846407    1.1621278    0.8731382
H    -4.4549772    1.4161217    0.9084980
H    -4.8598009   -2.3992986   -1.0384863
H    -2.3988359   -2.6452186   -1.0717401
H     4.7380462    1.6002164    0.0285317
O     5.7462895   -0.8695668    0.2493888
O    -6.1157204   -0.3920090   -0.1154217
H    -0.5833451   -3.0407839    0.0909975
H     6.3065964   -0.0751114    0.2472106
H    -6.4455770    0.4635881    0.2520571
O     0.6552756    4.0232506   -0.7707838
H     1.0112154    3.9310861   -1.6560571
H     0.3475682    3.1175978   -0.5531290
O     7.4894015    1.3418506    0.2993300
H     8.1799847    1.4185970   -0.3627000
H     7.8921171    1.5813055    1.1370310
```

```
O    -7.0327691    1.9394503    0.8407339
H    -7.6220047    2.0108135    1.5958292
H    -7.2230614    2.6865263    0.2674263

38
CT
O     1.3024510   -2.3378514    0.0998070
C    -0.0208605   -2.1246328   -0.0179418
C    -0.5577675   -0.8685821   -0.1194027
C     0.2671373    0.3143500   -0.1653598
O    -0.2236149    1.4785100   -0.3192829
C     1.6878623    0.0582436   -0.0499819
C     2.1663501   -1.2601891    0.0672669
C     2.6482972    1.0756443   -0.0627780
C     4.0121552    0.7939706    0.0391420
C     4.4438414   -0.5164410    0.1492994
C     3.5030835   -1.5619062    0.1634360
H     2.3082004    2.1009032   -0.1450747
H     3.8304379   -2.5895974    0.2514166
C    -2.0122866   -0.7572637   -0.0914325
C    -2.6311017    0.3921913    0.4776617
C    -3.9889084    0.5371071    0.4784044
C    -4.8023708   -0.4741437   -0.0777900
C    -4.2099464   -1.6418252   -0.6180691
C    -2.8511561   -1.7725591   -0.6196801
H    -1.9846402    1.1621275    0.8731379
H    -4.4549759    1.4161213    0.9084977
H    -4.8597996   -2.3992979   -1.0384861
H    -2.3988352   -2.6452179   -1.0717398
H     4.7380448    1.6002160    0.0285317
O     5.7462879   -0.8695665    0.2493888
O    -6.1157187   -0.3920089   -0.1154216
H    -0.5833450   -3.0407830    0.0909975
H     6.3065947   -0.0751113    0.2472105
H    -6.4455752    0.4635880    0.2520570
O     0.6552754    4.0232495   -0.7707836
H     1.0112151    3.9310850   -1.6560566
H     0.3475681    3.1175969   -0.5531289
O     7.4893994    1.3418502    0.2993300
H     8.1799824    1.4185966   -0.3626999
H     7.8921149    1.5813051    1.1370307
```

```
O    -7.0327672    1.9394497    0.8407337
H    -7.6220026    2.0108129    1.5958288
H    -7.2230594    2.6865256    0.2674262

38
LE
O     1.3268038   -2.2885055   -0.0247799
C    -0.0090566   -2.0515931   -0.0413408
C    -0.5691269   -0.8039078    0.0006870
C     0.3036425    0.3103196    0.1347250
O    -0.1181752    1.5492299    0.2220821
C     1.7186949    0.0844014    0.1498008
C     2.1895726   -1.2409499    0.0554463
C     2.6743268    1.1081599    0.2330229
C     4.0309766    0.8219285    0.2318003
C     4.4676260   -0.4979285    0.1478134
C     3.5334766   -1.5354001    0.0649201
H     2.3429738    2.1370778    0.2925638
H     3.8628325   -2.5646797    0.0065644
C    -2.0296877   -0.6521170   -0.0888127
C    -2.6941592    0.3827293    0.5666901
C    -4.0715428    0.4976610    0.5259002
C    -4.8326257   -0.4335585   -0.1889289
C    -4.1820341   -1.4588912   -0.8698523
C    -2.8042509   -1.5524867   -0.8298288
H    -2.1216197    1.1146912    1.1284443
H    -4.5683842    1.3157549    1.0366563
H    -4.7729348   -2.1605496   -1.4465775
H    -2.3150288   -2.3400652   -1.3919828
H     4.7574792    1.6246243    0.2927843
O     5.7689443   -0.8431782    0.1450127
O    -6.1728638   -0.3642197   -0.2740325
H    -0.5763699   -2.9685203   -0.0315697
H     6.3316110   -0.0521365    0.2278635
H    -6.5129243    0.3943757    0.2352767
O     0.7542798    3.4884531   -1.7303054
H     1.2846846    2.9104131   -2.2824373
H     0.2696998    2.8840655   -1.1511499
O     7.4634841    1.3541809    0.4153962
H     8.0995267    1.5497827   -0.2767227
H     7.9278779    1.4707905    1.2476482
```

O	−7.2261969	1.7807655	1.1488429
H	−7.7990496	1.6115634	1.9005723
H	−7.6230818	2.5106694	0.6673825

Table 2: Gas phase vertical excitation energies ΔE_{abs}(eV), oscillator strengths f and dipole moments μ (Debye) for daidzein at the PBE0/TZVP ground state geometry.

State	CC2	PBE0	CC2	PBE0	CC2	PBE0
	ΔE_{abs}		f		μ	
S_0					2.47	2.31
S_1	3.95	3.95	0.000	0.021	2.49	7.15
S_2	4.43	3.89	0.098	0.009	4.58	4.34
S_3	4.68	4.60	0.002	0.031	1.39	1.06
S_4	4.78	4.90	0.030	0.033	2.38	0.79
S_5	5.09	4.79	0.544	0.615	2.74	0.83